



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2019

Strong convergence of eigenangles and eigenvectors for the circular unitary ensemble

Maples, Kenneth ; Najnudel, Joseph ; Nikeghbali, Ashkan

Abstract: It is known that a unitary matrix can be decomposed into a product of complex reflections, one for each dimension, and that these reflections are independent and uniformly distributed on the space where they live if the initial matrix is Haar-distributed. If we take an infinite sequence of such reflections, and consider their successive products, then we get an infinite sequence of unitary matrices of increasing dimension, all of them following the circular unitary ensemble. In this coupling, we show that the eigenvalues of the matrices converge almost surely to the eigenvalues of the flow, which are distributed according to a sine-kernel point process, and we get some estimates of the rate of convergence. Moreover, we also prove that the eigenvectors of the matrices converge almost surely to vectors which are distributed as Gaussian random fields on a countable set.

DOI: <https://doi.org/10.1214/18-aop1311>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-177359>

Journal Article

Published Version

Originally published at:

Maples, Kenneth; Najnudel, Joseph; Nikeghbali, Ashkan (2019). Strong convergence of eigenangles and eigenvectors for the circular unitary ensemble. *The Annals of Probability*, 47(4):2417-2458.

DOI: <https://doi.org/10.1214/18-aop1311>

STRONG CONVERGENCE OF EIGENANGLES AND EIGENVECTORS FOR THE CIRCULAR UNITARY ENSEMBLE

BY KENNETH MAPLES^{*}, JOSEPH NAJNUDEL[†] AND ASHKAN NIKEGHBALI^{*}

Universität Zürich^{} and University of Cincinnati[†]*

It is known that a unitary matrix can be decomposed into a product of complex reflections, one for each dimension, and that these reflections are independent and uniformly distributed on the space where they live if the initial matrix is Haar-distributed. If we take an infinite sequence of such reflections, and consider their successive products, then we get an infinite sequence of unitary matrices of increasing dimension, all of them following the circular unitary ensemble.

In this coupling, we show that the eigenvalues of the matrices converge almost surely to the eigenvalues of the flow, which are distributed according to a sine-kernel point process, and we get some estimates of the rate of convergence. Moreover, we also prove that the eigenvectors of the matrices converge almost surely to vectors which are distributed as Gaussian random fields on a countable set.

Notation. If $v \in \mathbb{C}^n$ is a vector, then we write $v[m]$ for the image of v under the canonical projection map $\mathbb{C}^n \rightarrow \mathbb{C}^m$ onto the first m standard basis vectors and we write $(v)_k$ for its k th coordinate.

We write $O(n)$ for the orthogonal group of dimension n , that is, the group of invertible operators on \mathbb{R}^n which preserve the standard real inner product. We write $U(n)$ for the unitary group of dimension n which preserves the standard complex inner product. For any vector space V , real or complex, we write $GL(V)$ for the group of invertible transformations. We always write 1 for the identity operator in every space.

We write $\mathbb{U} = U(1)$ for the unit circle in \mathbb{C} , that is, those complex numbers with modulus 1.

Calligraphic characters denote σ -algebras, that is, $\mathcal{A}, \mathcal{B}, \mathcal{C}$, etc. If \mathcal{A} and \mathcal{B} are σ -algebras on a common set, then $\mathcal{A} \vee \mathcal{B}$ denotes the smallest σ -algebra containing both \mathcal{A} and \mathcal{B} .

We also write $a \vee b$, for $a, b \geq 0$, to mean $\max(a, b)$ and $a \wedge b$ to mean $\min(a, b)$.

We employ asymptotic notation for inequalities where precise constants are not important. In particular, we write $X = \mathcal{O}(Y)$ to mean that there exists a

Received May 2017; revised September 2018.

MSC2010 subject classifications. 60B20, 60F15.

Key words and phrases. Random matrix, circular unitary ensemble, convergence of eigenvalues, convergence of eigenvectors, virtual isometries, complex reflections.

constant $C > 0$ such that $|X| \leq CY$. All constants are assumed to be absolute unless indicated otherwise by an appropriate subscript; for example, we write $f_n(x) = \mathcal{O}_n(g_n(x))$ to mean that there are constants C_n , one for each value of n , such that $|f_n(x)| \leq C_n g_n(x)$ for all x . We also use (modified) Vinogradov notation, where $X \lesssim Y$ means $X = \mathcal{O}(Y)$, for convenience.

If t is a real number, we write $\lfloor t \rfloor$ its integer part.

If H is a Hilbert space with scalar product $\langle \cdot, \cdot \rangle$, and if $F \subset H$, then $F^\perp = \{x \in H; \langle x, y \rangle = 0 \forall y \in F\}$. If H is a complex Hilbert space, then we will always use the scalar product which is linear in the first variable and conjugate linear in the second, that is, $\langle ax, by \rangle = a\bar{b}\langle x, y \rangle$.

1. Introduction. It has been observed that for many models of random matrices, the eigenvalues have a limiting short-scale behavior when the dimension goes to infinity which depends on the global symmetries of the model, but not on its detailed features. For example, the Gaussian Orthogonal Ensemble (GOE), for which the matrices are real symmetric with independent Gaussian entries on and above the diagonal, corresponds to a limiting short-scale behavior for the eigenvalues that is also obtained for several other models of random real symmetric matrices. Similarly, the limiting spectral behavior of a large class of random hermitian and unitary ensembles, including the Gaussian Unitary Ensemble (GUE, with independent, complex Gaussians above the diagonal), and the Circular Unitary Ensemble (CUE, corresponding to the Haar measure on the unitary group of a given dimension), involves a remarkable random point process, called the *determinantal sine-kernel process*. It is a point process for which the k -point correlation function is given by

$$\rho_k(x_1, \dots, x_k) = \det \left(\frac{\sin(\pi(x_p - x_q))}{\pi(x_p - x_q)} \right)_{1 \leq p, q \leq k}.$$

From an observation of Montgomery in 1972, it has also been conjectured that the limiting short-scale behavior of the imaginary parts of the zeros of the Riemann zeta function is also described by a determinantal sine-kernel process. This similar behavior supports the conjecture of Hilbert and Pólya, who suggested that the non-trivial zeros of the Riemann zeta functions should be interpreted as the spectrum of an operator $\frac{1}{2} + iH$ with H an unbounded Hermitian operator.

In 1999, Katz and Sarnak [6] gave a proof of the Montgomery conjecture in the function field case. It appears from this work (which relies, among other things, on existing work developed for the proof of Weil's conjectures, most notably Deligne's equidistribution theorems) that the classical compact groups (e.g., the unitary group, the orthogonal group, the symplectic group, etc.) endowed with the Haar measure play a central role in the corresponding spectral interpretation. The regime studied by Katz and Sarnak (fixed genus, and number of elements of the base field going to infinity) does not contain arithmetic in the limiting spectral interpretation (this is an effect of equidistribution theorems when the number of

elements of the base field goes to infinity). On the other hand, the problem is still open in the *problematic regime* where the number of elements of the field is fixed and the genus goes to infinity. There it is not clear how random matrix statistics and arithmetic mix together in the limit (see [5] and [9] for more discussion on this particular aspect). A similar phenomenon occurs for the Riemann zeta function. Inspired by the work of Katz and Sarnak, Keating and Snaith [7] suggest to use the characteristic polynomial of random unitary matrices to model the distribution of values of the Riemann zeta function on the critical line and propose a conjecture for the moments of the Riemann zeta function on the critical line where again random matrix statistics and arithmetic mix in a mysterious way.

Following this body of work, several questions have naturally emerged. In particular, Katz and Sarnak asked whether it is possible to give a meaning to strong convergence (i.e., almost sure convergence) for the eigenvalues of random unitary matrices to the determinantal sine kernel point process. This problem was first solved by Borodin and Olshanski in [1] and then together with Paul Bourgade, the second and third authors of this paper proposed in [3] an alternative solution with the so called virtual isometries (see below for more details on the construction of virtual isometries). Our approach was probabilistic (we used coupling techniques ideas) and also quantitative: we were able to quantify the rate of convergence to the sine kernel point process. The goal of this paper is twofold:

- to improve on our estimates in the rate of convergence to the sine-kernel point process by refining several other estimates;
- to give a complete panorama of the spectral analysis of virtual isometries by establishing quantitative strong convergence for the eigenvectors as well.

We believe that these new and/or refined estimates can be very useful in tackling other problems at the interface of random matrix theory and analytic number theory. For instance, in the companion paper [4], we provide a new approach to ratios of characteristic polynomials and solve a conjecture on the limit of ratios of characteristic polynomials, where we use the estimates of this paper. In another companion paper, we use the spectral analysis of this paper to construct a flow of operators, constructed from virtual isometries, and whose spectrum is the sine-kernel point process.

We now briefly recall the notion of *virtual isometry*, which has been introduced in [3], generalizing both the notion of *virtual permutation* studied by Kerov, Olshanski and Vershik [8], and the previous notion of virtual unitary group introduced by Neretin [11]. A virtual isometry is a sequence of random unitary matrices $(u_n)_{n \geq 1}$ constructed in the following way:

1. One considers a sequence $(x_n)_{n \geq 1}$ of independent random vectors, x_n being uniform on the unit sphere of \mathbb{C}^n .
2. Almost surely, for all $n \geq 1$, x_n is different from the last basis vector e_n of \mathbb{C}^n , which implies that there exists a unique $r_n \in U(n)$ such that $r_n(e_n) = x_n$ and $r_n - I_n$ has rank one.

3. We define $(u_n)_{n \geq 1}$ by induction as follows: $u_1 = x_1$ and for all $n \geq 2$,

$$u_n = r_n \begin{pmatrix} u_{n-1} & 0 \\ 0 & 1 \end{pmatrix}.$$

It was proven in [2] that with this construction, u_n follows, for all $n \geq 1$, the Haar measure on $U(n)$. From now on, we always assume that the sequence $(u_n)_{n \geq 1}$ is defined with this coupling.

For each value of n , let $\lambda_1^{(n)}, \dots, \lambda_n^{(n)}$ be the eigenvalues of u_n , ordered counter-clockwise, starting from 1: they are almost surely pairwise distinct and different from 1. If $1 \leq k \leq n$, we denote by $\theta_k^{(n)}$ the argument of $\lambda_k^{(n)}$, taken in the interval $(0, 2\pi)$: $\theta_k^{(n)}$ is the k th strictly positive eigenangle of u_n . If we consider all the eigenangles of u_n , taken not only in $(0, 2\pi)$ but in the whole real line, we get a (2π) -periodic set with n points in each period. If the eigenangles are indexed increasingly by \mathbb{Z} , we obtain a sequence

$$\dots < \theta_{-1}^{(n)} < \theta_0^{(n)} < 0 < \theta_1^{(n)} < \theta_2^{(n)} < \dots,$$

for which $\theta_{k+n}^{(n)} = \theta_k^{(n)} + 2\pi$ for all $k \in \mathbb{Z}$.

It is also convenient to extend the sequence of eigenvalues as a n -periodic sequence indexed by \mathbb{Z} , in such a way that for all $k \in \mathbb{Z}$,

$$\lambda_k^{(n)} = \exp(i\theta_k^{(n)}).$$

Note that in the notion of virtual isometry defined here, the vectors of the canonical basis of \mathbb{C}^n play a particular role. One could attempt to generalize the notion of virtual isometries by considering sequences of unitary operators on E_n , $n \geq 1$, where $(E_n)_{n \geq 1}$ is a sequence of complex inner product spaces, E_n being of dimension n . However, this reduces to the particular case $E_n = \mathbb{C}^n$ by a change of basis and so we have chosen to use the standard basis for simplicity.

Next, for $1 \leq k \leq n$, let $f_k^{(n)} \in \mathbb{C}^n$ denote a unit length representative of the eigenspace for $\lambda_k^{(n)}$. Then if we expand $r_{n+1}(e_{n+1})$ in the basis of eigenvectors

$$r_{n+1}(e_{n+1}) = \sum_{j=1}^n \mu_j^{(n)} f_j^{(n)} + v_n e_{n+1}$$

then the eigenvalues of u_{n+1} are precisely the zeros of the rational equation

$$\sum_{j=1}^n |\mu_j^{(n)}|^2 \frac{\lambda_j^{(n)}}{\lambda_j^{(n)} - z} + \frac{|1 - v_n|^2}{1 - z} = 1 - \bar{v}_n$$

and the eigenvectors of u_{n+1} are given by the $n + 1$ equations

$$C_k f_k^{(n+1)} = \sum_{j=1}^n \frac{\mu_j^{(n)}}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} f_j^{(n)} + \frac{v_n - 1}{1 - \lambda_k^{(n+1)}} e_{n+1}$$

for $1 \leq k \leq n+1$; here, $C_k \in \mathbb{R}^+$ is a constant so that $f_k^{(n+1)}$ has unit length. As above, we will also extend, when needed, the notation $f_k^{(n)}$, $\mu_j^{(n)}$, in such a way that the sequences $(f_k^{(n)})_{k \in \mathbb{Z}}$ and $(\mu_j^{(n)})_{j \in \mathbb{Z}}$ are n -periodic.

In [3], it is shown that if $(u_n)_{n \geq 1}$ follows this distribution, then for all k , the k th positive (resp., negative) eigenangle $\theta_k^{(n)}$ (resp. $\theta_{1-k}^{(n)}$) of u_n , multiplied by $n/2\pi$ (i.e., the inverse of the average spacing between eigenangles for any matrix in $U(n)$), converges almost surely to a random variable y_k (resp., y_{1-k}). The random set $(y_k)_{k \in \mathbb{Z}}$ is a determinantal sine-kernel process, and for each k , the convergence holds with a rate dominated by some negative power of n . In the present paper, we improve our estimate of this rate, and more importantly, we prove that almost sure convergence not only holds for the eigenangles of u_n , but also for the components of the corresponding eigenvectors. More precisely, we show that, for all $k, \ell \geq 1$, the ℓ th component of the eigenvector of u_n associated to the k th positive (resp., negative) eigenangle converges almost surely to a nonzero limit when n goes to infinity, if the norm of the eigenvector is taken equal to \sqrt{n} and if the phases are suitably chosen. Note that taking a norm equal to \sqrt{n} is natural in this setting: with this normalization, the expectation of the squared modulus of each coordinate of a given eigenvector of u_n is equal to 1, so we can expect a convergence to a nontrivial limit. If the norm of the eigenvectors is taken equal to 1 instead of \sqrt{n} , then the coordinates converge to zero when n goes to infinity.

The precise statement of our main results are given as follows.

THEOREM 1.1. *With the notation above, the following estimate holds almost surely:*

$$\frac{n}{2\pi} \theta_k^{(n)} = y_k + \mathcal{O}((1+k^2)n^{-\frac{1}{3}+\varepsilon}),$$

for all $n \geq 1$, $|k| \leq n^{1/4}$ and $\varepsilon > 0$, where the implied constant may depend on $(u_m)_{m \geq 1}$ and ε , but not on n and k .

THEOREM 1.2. *Let $(u_n)_{n \geq 1}$ be a virtual isometry, following the Haar measure. For $k \in \mathbb{Z}$ and $n \geq 1$, let $v_k^{(n)}$ be a unit eigenvector corresponding to the k th smallest nonnegative eigenangle of u_n for $k \geq 1$, and the $(1-k)$ th largest strictly negative eigenangle of u_n for $k \leq 0$. Then for all $k \in \mathbb{Z}$, there almost surely exist some complex numbers $(\psi_k^{(n)})_{n \geq 1}$ of modulus 1, and a sequence $(t_{k,\ell})_{\ell \geq 1}$, such that for all $\ell \geq 1$,*

$$\sqrt{n} \langle \psi_k^{(n)} v_k^{(n)}, e_\ell \rangle \xrightarrow{n \rightarrow \infty} t_{k,\ell}.$$

Almost surely, for all $k \in \mathbb{Z}$, the sequence $(t_{k,\ell})_{\ell \geq 1}$ depends, up to a multiplicative factor of modulus one, only on the virtual rotation $(u_n)_{n \geq 1}$. Moreover, if $(\psi_k)_{k \in \mathbb{Z}}$ is a sequence of i.i.d., uniform variables on \mathbb{U} , independent of $(t_{k,\ell})_{\ell \geq 1}$, then $(\psi_k t_{k,\ell})_{k \in \mathbb{Z}, \ell \geq 1}$ is an i.i.d. family of standard complex Gaussian variables ($\mathbb{E}[|\psi_k t_{k,\ell}|^2] = 1$).

REMARK 1.3. The vectors $v_k^{(n)}$ are equal to $f_k^{(n)}$, up to a multiplicative factor of modulus 1. The independent phases ψ_k introduced in the last part of the theorem are needed in order to get i.i.d. complex Gaussian variables. This is not the case, for example, if we normalize $(t_{k,\ell})_{\ell \geq 1}$ in such a way that $t_{k,1} \in \mathbb{R}_+$.

In Section 3, we make the following key observation: the sequence of eigenvalues $\lambda_k^{(n)}$, with $1 \leq k \leq n$ and $n \geq 1$, is independent of the argument of the coefficients $\mu_j^{(n)}$, with $1 \leq j \leq n$ and $n \geq 1$. Therefore, we can consider the sequence of eigenvalues of the virtual isometry and prove that it converges almost surely, and then, *conditioning on the eigenvalues of every matrix in the virtual isometry*, consider the sequence of eigenvectors and show that they also converge in a suitable sense.

The first part of this plan is carried out in Section 4, where Theorem 1.1 is proven.

Next, in Section 5 we condition on the eigenangles of the entire sequence of matrices. We show that for each fixed k there is a renormalization factor $D_k^{(n)}$ so that for each $\ell \geq 1$ the sequence $\langle D_k^{(n)} f_k^{(n)}, e_\ell \rangle$ is a martingale which converges in L^2 and almost surely to a limiting value $g_{k,\ell}$. In Section 6, Theorem 1.2 is deduced from this convergence.

2. Spectral analysis of the virtual isometries. It is classical that if u_n is a random unitary matrix following the Haar measure on $U(n)$, then the distribution of the eigenangles of u_n , multiplied by $n/2\pi$, converges in law to a determinantal sine-kernel process. In fact, this result can be found in the literature under the form of the convergence of the correlation functions against a suitably chosen family of test functions. However, we were not able to find a statement with its proof on the fact that the convergence takes place in the sense of weak convergence of point processes, using Laplace functionals. So for completeness, we give such a statement below and postpone its proof until Appendix B.

PROPOSITION 2.1. *Let E_n denote the set of eigenvalues taken in $(-\pi, \pi]$ and multiplied by $n/2\pi$ of a random unitary matrix of size n following the Haar measure. Let us also define for $y \neq y'$,*

$$K^{(\infty)}(y, y') = \frac{\sin[\pi(y' - y)]}{\pi(y' - y)}$$

and

$$K^{(\infty)}(y, y) = 1.$$

Then there exists a point process E_∞ such that for all $r \geq 1$, and for all Borel measurable and bounded functions F with compact support from \mathbb{R}^r to \mathbb{R} , we

have

$$\mathbb{E} \left(\sum_{x_1 \neq \dots \neq x_r \in E_\infty} F(x_1, \dots, x_r) \right) = \int_{\mathbb{R}^r} F(y_1, \dots, y_r) \rho_r^{(\infty)}(y_1, \dots, y_r) dy_1 \cdots dy_r,$$

where

$$\rho_r^{(\infty)}(y_1, \dots, y_r) = \det((K^{(\infty)}(y_j, y_k))_{1 \leq j, k \leq r}).$$

Moreover, the point process E_n converges to E_∞ in the following sense: for all Borel measurable bounded functions f with compact support from \mathbb{R} to \mathbb{R} ,

$$\sum_{x \in E_n} f(x) \xrightarrow{n \rightarrow \infty} \sum_{x \in E_\infty} f(x),$$

where the convergence above holds in law.

In [3], it was proven that the eigenangles of a virtual isometry, taken according to Haar measure and renormalizing the eigenangles by the dimension n , converge almost surely to a point process with this determinantal distribution. Precisely, we have the following.

PROPOSITION 2.2. *Let $(u_n)_{n \geq 1}$ be a random virtual isometry following the Haar measure. The eigenvalues of u_n are almost surely distinct and different from 1, and then, as explained before, it is possible to order the eigenangles as an increasing sequence indexed by \mathbb{Z} :*

$$\dots < \theta_{-2}^{(n)} < \theta_{-1}^{(n)} < \theta_0^{(n)} < 0 < \theta_1^{(n)} < \theta_2^{(n)} < \dots.$$

Moreover, almost surely, for all $m \in \mathbb{Z}$, there exists y_m such that

$$\frac{n}{2\pi} \theta_m^{(n)} = y_m + \mathcal{O}(n^{-\varepsilon}),$$

when n goes to infinity, $\varepsilon > 0$ being some universal constant, and the process $(y_m)_{m \in \mathbb{Z}}$ is a determinantal sine-kernel process.

REMARK 2.3. In this proposition, the periodic extension of the sequence $(\theta_k^{(n)})_{1 \leq k \leq n}$ is needed to define y_m for nonpositive values of m . We also note that y_m depends only on the behavior of $\theta_m^{(n)}$ for a fixed value of m . Informally, the limiting sine-kernel process $(y_m)_{m \in \mathbb{Z}}$ depends only on the behavior of the eigenvalues of u_n which are close to 1.

We want to understand the behavior of the eigenvectors of u_n as n goes to infinity. Our method will give a formula for the spectrum and eigenvectors of u_{n+1} in terms of the spectrum and eigenvectors of u_n .

We assume throughout that for each $n \geq 1$, the n eigenvalues of u_n are distinct; this holds almost surely for virtual isometries constructed according to the Haar measure.

We recall that the eigenvalues of u_n , $\lambda_1^{(n)}, \lambda_2^{(n)}, \dots, \lambda_n^{(n)}$, are ordered in such a way that $\lambda_k^{(n)} = e^{i\theta_k^{(n)}}$, and

$$0 < \theta_1^{(n)} < \dots < \theta_n^{(n)} < 2\pi.$$

Moreover, the eigenangles enjoy a property of periodicity: for all $k \in \mathbb{Z}$, $\theta_{k+n}^{(n)} = \theta_k^{(n)} + 2\pi$.

As all the eigenvalues are distinct, each eigenvalue corresponds to a one-dimensional eigenspace. We can therefore write $f_1^{(n)}, \dots, f_n^{(n)}$ for the family of unit length eigenvectors of u_n , which are well defined up to a complex phase: the notation $f_k^{(n)}$ is then extended n -periodically to all $k \in \mathbb{Z}$.

Let $x_n = u_n(e_n)$ and let r_n denote the unique reflection on \mathbb{C}^n mapping e_n to x_n . Therefore, we have $u_{n+1} = r_{n+1} \circ (u_n \oplus 1)$. It is natural to decompose x_{n+1} into the basis given by $\iota(f_1^{(n)}), \dots, \iota(f_n^{(n)}), e_{n+1}$, where $\iota: \mathbb{C}^n \rightarrow \mathbb{C}^{n+1}$ is the inclusion which maps (x_1, \dots, x_n) to $(x_1, \dots, x_n, 0)$. Identifying $f_k^{(n)}$ and $\iota(f_k^{(n)})$, we then have

$$x_{n+1} = \sum_{k=1}^n \mu_k^{(n)} f_k^{(n)} + v_n e_{n+1}$$

for some $\mu_k^{(n)}$ ($1 \leq k \leq n$) and v_n such that $|\mu_1^{(n)}|^2 + \dots + |\mu_n^{(n)}|^2 + |v_n|^2 = 1$. Again, it can be convenient to consider $\mu_k^{(n)}$ for all $k \in \mathbb{Z}$, by a n -periodic extension of the sequence. The following result gives the spectral decomposition of u_{n+1} in function of the decomposition of u_n and x_{n+1} .

THEOREM 2.4 (Spectral decomposition). *On the event that the coefficients $\mu_1^{(n)}, \dots, \mu_n^{(n)}$ are all different from zero and that the n eigenvalues of u_n are all distinct (which holds almost surely under the uniform measure on U^∞), the eigenvalues of u_{n+1} are the unique roots of the rational equation*

$$\sum_{j=1}^n |\mu_j^{(n)}|^2 \frac{\lambda_j^{(n)}}{\lambda_j^{(n)} - z} + \frac{|1 - v_n|^2}{1 - z} = 1 - \bar{v}_n$$

on the unit circle. Furthermore, they interlace between 1 and the eigenvalues of u_n

$$0 < \theta_1^{(n+1)} < \theta_1^{(n)} < \theta_2^{(n+1)} < \dots < \theta_n^{(n)} < \theta_{n+1}^{(n+1)} < 2\pi,$$

and it is possible to choose the unit eigenvectors $f_k^{(n)}$ so that they satisfy the rela-

tion

$$(h_k^{(n+1)})^{\frac{1}{2}} f_k^{(n+1)} = \sum_{j=1}^n \frac{\mu_j^{(n)}}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} f_j^{(n)} + \frac{v_n - 1}{1 - \lambda_k^{(n+1)}} e_{n+1},$$

where $h_k^{(n+1)}$ is real and strictly positive.

PROOF. Let f be an eigenvector of u_{n+1} with corresponding eigenvalue z . Then we have

$$f = \sum_{j=1}^n a_j f_j^{(n)} + b e_{n+1},$$

where a_1, \dots, a_n, b are (as yet unknown) complex numbers, not all zero. Our goal is to write these coefficients in terms of x_{n+1} and the eigenvalues of u_n .

We have

$$\begin{aligned} z f &= u_{n+1} f \\ &= u_{n+1} \left(\sum_{j=1}^n a_j f_j^{(n)} + b e_{n+1} \right) \\ &= \sum_{j=1}^n a_j u_{n+1} f_j^{(n)} + b u_{n+1} e_{n+1} \\ &= \sum_{j=1}^n a_j \lambda_j^{(n)} r_{n+1} f_j^{(n)} + b x_{n+1}. \end{aligned}$$

We recall that for all $t \in \mathbb{C}^{n+1}$, $r_{n+1}(t)$ is given by

$$r_{n+1}(t) = t + \frac{\langle t, x_{n+1} - e_{n+1} \rangle}{\langle e_{n+1}, x_{n+1} - e_{n+1} \rangle} (x_{n+1} - e_{n+1})$$

so that

$$z f = \sum_{j=1}^n a_j \lambda_j^{(n)} \left(f_j^{(n)} + \frac{\langle f_j^{(n)}, x_{n+1} - e_{n+1} \rangle}{\langle e_{n+1}, x_{n+1} - e_{n+1} \rangle} (x_{n+1} - e_{n+1}) \right) + b x_{n+1}.$$

Now we decompose

$$x_{n+1} = \sum_{k=1}^n \mu_k^{(n)} f_k^{(n)} + v_n e_{n+1}$$

and

$$\begin{aligned} zf &= \sum_{j=1}^n a_j \lambda_j^{(n)} \left(f_j^{(n)} + \frac{\overline{\mu_j^{(n)}}}{v_n - 1} (x_{n+1} - e_{n+1}) \right) + bx_{n+1} \\ &= \sum_{j=1}^n a_j \lambda_j^{(n)} f_j^{(n)} + \left(\sum_{\ell=1}^n a_\ell \lambda_\ell^{(n)} \frac{\overline{\mu_\ell^{(n)}}}{v_n - 1} \right) (x_{n+1} - e_{n+1}) + bx_{n+1}. \end{aligned}$$

Because $f_1^{(n)}, \dots, f_n^{(n)}, e_{n+1}$ is a basis for \mathbb{C}^{n+1} , we deduce the system of $n+1$ equations

$$za_j = a_j \lambda_j^{(n)} + \mu_j^{(n)} \sum_{\ell=1}^n a_\ell \lambda_\ell^{(n)} \frac{\overline{\mu_\ell^{(n)}}}{v_n - 1} + b \mu_j^{(n)}$$

for $j = 1, \dots, n$ and

$$zb = b + (v_n - 1) \sum_{\ell=1}^n a_\ell \lambda_\ell^{(n)} \frac{\overline{\mu_\ell^{(n)}}}{v_n - 1} + b(v_n - 1).$$

For $z \notin \{\lambda_1^{(n)}, \dots, \lambda_n^{(n)}, 1\}$, let us consider the linear transform $Q : \mathbb{C}^{n+1} \rightarrow \mathbb{C}^{n+1}$ whose matrix representation in the basis $f_1^{(n)}, \dots, f_n^{(n)}, e_{n+1}$ is

$$Q = I + wv^t,$$

where

$$w = \begin{pmatrix} \frac{\mu_1^{(n)}}{\lambda_1^{(n)} - z} \\ \vdots \\ \frac{\mu_n^{(n)}}{\lambda_n^{(n)} - z} \\ \frac{v_n - 1}{1 - z} \end{pmatrix}$$

and

$$v^t = \left(\lambda_1^{(n)} \frac{\overline{\mu_1^{(n)}}}{v_n - 1}, \dots, \lambda_n^{(n)} \frac{\overline{\mu_n^{(n)}}}{v_n - 1}, 1 \right).$$

Then, the above system can be written

$$Qf = 0.$$

Clearly, $\text{rank } Q \in \{n, n+1\}$. If it has full rank then $f = 0$, but we assume a priori that z is an eigenvalue for u_{n+1} and so has a nontrivial eigenspace. Thus we must have $\text{rank } Q = n$ and

$$0 = Qf = f + w(v^t f).$$

The right-hand side can only vanish if f is proportional to w , so $f = \alpha w$ for some complex constant $\alpha \in \mathbb{C} \setminus \{0\}$ and $v^t w = -1$. In particular,

$$\sum_{j=1}^n \frac{\lambda_j^{(n)} |\mu_j^{(n)}|^2}{\lambda_j^{(n)} - z} + \frac{|v_n - 1|^2}{1 - z} = 1 - \overline{v_n},$$

as required.

Conversely, if $z \notin \{\lambda_1^{(n)}, \dots, \lambda_n^{(n)}, 1\}$ satisfies this equation, then

$$Qw = w + w(v^t w) = w + w(-1) = 0,$$

which implies that w is an eigenvector of u_{n+1} for the eigenvalue z .

Let us now show that the eigenvalues $z \notin \{\lambda_1^{(n)}, \dots, \lambda_n^{(n)}, 1\}$ of u_{n+1} strictly interlace between 1 and the eigenvalues of u_n : since u_{n+1} has at most $n + 1$ eigenvalues, this implies that $\lambda_1^{(n)}, \dots, \lambda_n^{(n)}, 1$ are not eigenvalues of u_{n+1} .

Define the rational function $\Phi : S^1 \rightarrow \mathbb{C} \cup \{\infty\}$ by

$$\Phi(z) = \sum_{j=1}^n \frac{\lambda_j^{(n)} |\mu_j^{(n)}|^2}{\lambda_j^{(n)} - z} + \frac{|v_n - 1|^2}{1 - z} - (1 - \overline{v_n}).$$

Note that Φ vanishes precisely on the eigenvalues of u_{n+1} which are different from $\lambda_1^{(n)}, \dots, \lambda_n^{(n)}, 1$. Recalling that $|\mu_1^{(n)}|^2 + \dots + |\mu_n^{(n)}|^2 + |v_n|^2 = 1$, we can rearrange the expression of Φ to the equivalent form

$$\Phi(z) = \frac{1}{2} \left(\sum_{j=1}^n |\mu_j^{(n)}|^2 \frac{\lambda_j^{(n)} + z}{\lambda_j^{(n)} - z} + |1 - v_n|^2 \frac{1 + z}{1 - z} - v_n + \overline{v_n} \right).$$

Hence, Φ takes values only in $i\mathbb{R} \cup \{\infty\}$, since for all $z \neq z' \in S^1$, $(z + z')/(z - z')$ is purely imaginary (the triangle $(-z', z, z')$ has a right angle at z). Note that for $z \in \{\lambda_1^{(n)}, \dots, \lambda_n^{(n)}, 1\}$, a unique term of the sum defining Φ is infinite, since by assumption, $\mu_1^{(n)}, \dots, \mu_n^{(n)}, 1 - v_n$ are nonzero and $\lambda_1^{(n)}, \dots, \lambda_n^{(n)}, 1$ are distinct: $\Phi(z) = \infty$.

Next, we consider $t \mapsto \Phi(e^{it})$ in a short interval $(\theta_j^{(n)} - \delta, \theta_j^{(n)} + \delta)$. Then, for $t = \theta_j^{(n)} + u$ in this interval,

$$\frac{\lambda_j^{(n)} + \lambda_j^{(n)} e^{iu}}{\lambda_j^{(n)} - \lambda_j^{(n)} e^{iu}} = \frac{1 + e^{iu}}{1 - e^{iu}} = 2iu^{-1} + \mathcal{O}(1)$$

while the other terms in $\Phi(e^{it})$ are uniformly bounded as $\delta \rightarrow 0$; likewise for the interval $(-\delta, \delta)$. In particular, $\Phi \rightarrow i\infty$ as $u \rightarrow 0$ from the right and $\Phi \rightarrow -i\infty$ as $u \rightarrow 0$ from the left. We therefore conclude, as Φ is continuous, that on each interval of the partition

$$(0, \theta_1^{(n)}) \cup (\theta_1^{(n)}, \theta_2^{(n)}) \cup \dots \cup (\theta_n^{(n)}, 2\pi)$$

of the unit circle into $n + 1$ parts, $t \mapsto \Phi(e^{it})$ must assume every value on the line $i\mathbb{R}$, and in particular must have at least one root. But we know that Φ has only $n + 1$ roots on the circle so there must be exactly one root in each part of the partition, which proves the interlacing property.

It remains to check the expression of the eigenvectors $(f_k^{(n+1)})_{1 \leq k \leq n+1}$ given in the theorem, but this expression is immediately deduced from the expression of the vector w involved in the operator Q defined above, and the fact that $\|f_k^{(n+1)}\| = 1$. \square

3. A filtration adapted to the virtual isometry. In our proof of convergence of eigenvalues and eigenvectors, we will use some martingale arguments. That is why in this section, we introduce a filtration related to our random virtual isometry. For $n \geq 1$, we define the σ -algebra $\mathcal{A}_n = \sigma\{\lambda_j^{(m)} \mid 1 \leq m \leq n, 1 \leq j \leq m\}$ and its limit $\mathcal{A} = \bigvee_{n=1}^{\infty} \mathcal{A}_n$.

LEMMA 3.1. *For all $n \geq 1$, the σ -algebra \mathcal{A}_n is equal, up to completion, to the σ -algebra generated by u_1 the variables $|\mu_j^{(m)}|$ and v_m for $1 \leq m \leq n - 1$ and $1 \leq j \leq m$.*

PROOF. By Theorem 2.4, the eigenvalues of u_{n+1} are almost surely the roots of the equation

$$\sum_{j=1}^n |\mu_j^{(n)}|^2 \frac{\lambda_j^{(n)}}{\lambda_j^{(n)} - z} + \frac{|1 - v_n|^2}{1 - z} = 1 - \bar{v}_n.$$

This equation depends only on $|\mu_j^{(n)}|$, $\lambda_j^{(n)}$ for $j = 1, \dots, n$, and v_n . By induction, we deduce that $\lambda_j^{(n+1)}$ is a measurable function of u_1 , $\{|\mu_j^{(m)}|\}_{1 \leq j \leq m, 1 \leq m \leq n}$ and $\{v_m\}_{1 \leq m \leq n}$.

Conversely, the above equation with z specialized to $\lambda_1^{(n+1)}, \dots, \lambda_{n+1}^{(n+1)}$ can be written in the form

$$Rv = w.$$

Here, w is a column vector of 1s, v is the column vector with entries

$$v^t = \left(\frac{|\mu_1^{(n)}|^2}{1 - \bar{v}_n}, \dots, \frac{|\mu_n^{(n)}|^2}{1 - \bar{v}_n}, 1 - v_n \right)$$

and R is an $(n + 1) \times (n + 1)$ matrix with entries

$$R_{k,j} = \frac{\lambda_j^{(n)}}{\lambda_j^{(n)} - \lambda_k^{(n+1)}}$$

for $1 \leq j \leq n$ and

$$R_{k,n+1} = \frac{1}{1 - \lambda_k^{(n+1)}}.$$

If we write $\tilde{R} = RS$, where S is the diagonal matrix with entries $S_{jj} = (\mu_j^{(n)} \lambda_j^{(n)})^{-1}$ ($1 \leq j \leq n$) and $S_{n+1,n+1} = v_n - 1$, then we see that the rows of \tilde{R} are the representations of the eigenvectors $f_1^{(n+1)}, \dots, f_{n+1}^{(n+1)}$ in the basis $f_1^{(n)}, \dots, f_n^{(n)}, e_{n+1}$ up to constants and, therefore, orthogonal. We conclude that \tilde{R} , and thus R , are invertible, so $Rv = w$ has a unique solution, which can be written in terms of the eigenvalues $\lambda_j^{(m)}$ for $1 \leq j \leq m$ and $m \leq n+1$. Thus we conclude that $|\mu_j^{(n)}|$ and v_n are measurable functions of $\lambda_j^{(m)}$ for $1 \leq j \leq m$ and $m \leq n+1$ as was to be shown. \square

For $1 \leq j \leq n$, we define the phase $\phi_j^{(n)}$ by $\mu_j^{(n)} = \phi_j^{(n)} |\mu_j^{(n)}|$, and the σ -algebras $\mathcal{B}_n = \mathcal{A} \vee \sigma\{\phi_j^{(m)} \mid 1 \leq m \leq n-1, 1 \leq j \leq m\}$ and $\mathcal{B} = \bigvee_{n=1}^{\infty} \mathcal{B}_n$.

LEMMA 3.2. *The σ -algebra \mathcal{B}_n is equal, up to completion, to the σ -algebra generated by \mathcal{A} and the eigenvectors $f_j^{(m)}$ for $1 \leq j \leq m$ and $1 \leq m \leq n$.*

PROOF. Again by Theorem 2.4, we can write the eigenvectors of u_{n+1} as functions of $\mu_j^{(n)} = \phi_j^{(n)} |\mu_j^{(n)}|$ ($1 \leq j \leq n$), v_n , $\lambda_j^{(m)}$ ($1 \leq j \leq m$, $m \in \{n, n+1\}$), and the eigenvectors of u_n . Clearly, $|\mu_j^{(n)}|$, v_n , and $\lambda_j^{(m)}$ for $m \in \{n, n+1\}$ are \mathcal{A} -measurable and $\phi_j^{(n)}$ is \mathcal{B}_{n+1} -measurable.

Conversely, we write

$$\langle (h_k^{(n+1)})^{1/2} f_k^{(n+1)}, f_j^{(n)} \rangle = \frac{\phi_j^{(n)} |\mu_j^{(n)}|}{\lambda_j^{(n)} - \lambda_k^{(n+1)}}$$

to find each $\phi_j^{(n)}$ as a function of the other variables. \square

4. Convergence of the eigenangles. In order to prove the convergence of the normalized eigenangles of u_n when n goes to infinity, we need the following lemma.

LEMMA 4.1. *Let $\varepsilon > 0$. Then, almost surely under the Haar measure on U^∞ , for $n \geq 1$ and $0 < k \leq n^{1/4}$, we have*

$$\frac{\theta_k^{(n+1)} |\mu_k^{(n)}|^2}{\theta_k^{(n)} - \theta_k^{(n+1)}} = 1 + \mathcal{O}(kn^{-\frac{1}{3} + \varepsilon})$$

and for $n \geq 1$ and $-n^{1/4} \leq k \leq 0$,

$$\frac{\theta_k^{(n+1)} |\mu_k^{(n)}|^2}{\theta_k^{(n)} - \theta_k^{(n+1)}} = \frac{\theta_k^{(n+1)} |\mu_{k+n}^{(n)}|^2}{\theta_k^{(n)} - \theta_k^{(n+1)}} = 1 + \mathcal{O}((1 + |k|)n^{-\frac{1}{3} + \varepsilon}).$$

REMARK 4.2. The implied constant in the $\mathcal{O}(\cdot)$ notation depends on $(u_m)_{m \geq 1}$ and ε : in particular, it is a random variable. However, for given $(u_m)_{m \geq 1}$ and ε , it does not depend on k and n .

PROOF. By symmetry of the situation, we can assume $k > 0$. Moreover, let us fix $\varepsilon \in (0, 0.01)$. We will suppose that the event $E := E_0 \cap E_1 \cap E_2 \cap E_3$ holds, where

$$E_0 = \{\theta_0^{(1)} \neq 0\} \cap \{\forall n \geq 1, v_n \neq 0\} \cap \{\forall n \geq 1, 1 \leq k \leq n, \mu_k^{(n)} \neq 0\},$$

$$E_1 = \{\exists n_0 \geq 1, \forall n \geq n_0, |v_n| \leq n^{-\frac{1}{2} + \varepsilon}\},$$

$$E_2 = \{\exists n_0 \geq 1, \forall n \geq n_0, 1 \leq k \leq n, |\mu_k^{(n)}| \leq n^{-\frac{1}{2} + \varepsilon}\},$$

$$E_3 = \{\exists n_0 \geq 1, \forall n \geq n_0, k \geq 1, n^{-\frac{5}{3} - \varepsilon} \leq \theta_{k+1}^{(n)} - \theta_k^{(n)} \leq n^{-1 + \varepsilon}\}.$$

It is possible to do this assumption, since by the result proven in Appendix A of the present paper, the event E occurs almost surely. As we will see now, this *a priori* information on the distribution of the eigenvalues of the random virtual isometry implies strong quantitative bounds on the change in eigenvalues of successive unitary matrices.

Recall from Theorem 2.4 that

$$\sum_{j=1}^n \frac{\lambda_j^{(n)} |\mu_j^{(n)}|^2}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} + \frac{|1 - v_n|^2}{1 - \lambda_k^{(n+1)}} = 1 - \bar{v}_n.$$

By using the n -periodicity of $\lambda_j^{(n)}$, $\mu_j^{(n)}$, $f_j^{(n)}$ with respect to j , we can write

$$(1) \quad \sum_{j \in J} \frac{\lambda_j^{(n)} |\mu_j^{(n)}|^2}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} + \frac{|1 - v_n|^2}{1 - \lambda_k^{(n+1)}} = 1 - \bar{v}_n,$$

where J is the random set of n consecutive integers, such that $\theta_k^{(n+1)} - \pi < \theta_j^{(n)} \leq \theta_k^{(n+1)} + \pi$. Iterating the lower bound on the distance between adjacent eigenvalues, given by the definition of the event E_3 , we get, for $j \in J \setminus \{k-1, k\}$,

$$|\theta_j^{(n)} - \theta_k^{(n+1)}| \gtrsim |k - j| n^{-\frac{5}{3} - \varepsilon},$$

and then

$$|\lambda_j^{(n)} - \lambda_k^{(n+1)}| \gtrsim |k - j| n^{-\frac{5}{3} - \varepsilon},$$

since $|\theta_j^{(n)} - \theta_k^{(n+1)}| \leq \pi$.

Likewise, we have by E_3 , $1 - \lambda_k^{(n+1)} = \mathcal{O}(kn^{-1+\varepsilon})$, and by E_2 , $|\mu_j^{(n)}|^2 = \mathcal{O}(n^{-1+2\varepsilon})$, which gives, for $j \in J \setminus \{k-1, k\}$,

$$\frac{\lambda_j^{(n)}(1 - \lambda_k^{(n+1)})|\mu_j^{(n)}|^2}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} \lesssim \frac{k}{|k-j|} n^{-\frac{1}{3}+4\varepsilon}.$$

Summing for j in $J \setminus \{k-1, k\}$, which is included in the interval $[k-1-n, k+n]$, gives

$$\sum_{j \in J \setminus \{k-1, k\}} \frac{\lambda_j^{(n)}(1 - \lambda_k^{(n+1)})|\mu_j^{(n)}|^2}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} = \mathcal{O}(kn^{-\frac{1}{3}+4\varepsilon} \log n) = \mathcal{O}(kn^{-\frac{1}{3}+5\varepsilon}).$$

Now, subtracting this equation from the product of (1) by $1 - \lambda_k^{(n+1)}$, and bounding $v_n = \mathcal{O}(n^{-\frac{1}{2}+\varepsilon})$ (by the property E_1) gives us the resulting equation

$$\begin{aligned} & \frac{\lambda_k^{(n)}(1 - \lambda_k^{(n+1)})|\mu_k^{(n)}|^2}{\lambda_k^{(n)} - \lambda_k^{(n+1)}} \mathbf{1}_{k \in J} + \frac{\lambda_{k-1}^{(n)}(1 - \lambda_k^{(n+1)})|\mu_{k-1}^{(n)}|^2}{\lambda_{k-1}^{(n)} - \lambda_k^{(n+1)}} \mathbf{1}_{k-1 \in J} \\ &= -1 + \mathcal{O}(kn^{-\frac{1}{3}+5\varepsilon}). \end{aligned}$$

Next, we estimate the first two terms in terms of the eigenangles. We find

$$1 - \lambda_k^{(n+1)} = -i\theta_k^{(n+1)} + \mathcal{O}((\theta_k^{(n+1)})^2)$$

and

$$\lambda_j^{(n)} - \lambda_k^{(n+1)} = i(\theta_j^{(n)} - \theta_k^{(n+1)})\lambda_j^{(n)} + \mathcal{O}((\theta_j^{(n)} - \theta_k^{(n+1)})^2)$$

for $j = k-1, k$. Collecting terms and using the trivial bounds give

$$\begin{aligned} (2) \quad & \frac{\theta_k^{(n+1)}|\mu_k^{(n)}|^2}{\theta_k^{(n)} - \theta_k^{(n+1)}} (1 + \mathcal{O}(kn^{-1+\varepsilon})) \mathbf{1}_{k \in J} \\ & + \frac{\theta_k^{(n+1)}|\mu_{k-1}^{(n)}|^2}{\theta_{k-1}^{(n)} - \theta_k^{(n+1)}} (1 + \mathcal{O}(kn^{-1+\varepsilon})) \mathbf{1}_{k-1 \in J} = 1 + \mathcal{O}(kn^{-\frac{1}{3}+5\varepsilon}). \end{aligned}$$

From Theorem 2.4, the eigenvalues of u_n and u_{n+1} interlace, so for n sufficiently large the real part of the first term is positive and the real part of the second term is negative. The real part of the right-hand side tends to 1 as n grows with k fixed, so the first term has real part bounded below for n sufficiently large. In particular,

$$\frac{\theta_k^{(n+1)}|\mu_k^{(n)}|^2}{\theta_k^{(n)} - \theta_k^{(n+1)}} \gtrsim 1.$$

Using the a priori bounds for $\theta_k^{(n+1)}$ and $|\mu_k^{(n)}|^2$, we find

$$\theta_k^{(n)} - \theta_k^{(n+1)} \lesssim kn^{-2+3\varepsilon}.$$

Hence,

$$\begin{aligned} \theta_k^{(n+1)} - \theta_{k-1}^{(n)} &= (\theta_k^{(n)} - \theta_{k-1}^{(n)}) - (\theta_k^{(n)} - \theta_k^{(n+1)}) \\ &\gtrsim n^{-\frac{5}{3}-\varepsilon} - \mathcal{O}(kn^{-2+3\varepsilon}) \gtrsim n^{-\frac{5}{3}-\varepsilon}, \end{aligned}$$

since $kn^{-2+3\varepsilon}/n^{-\frac{5}{3}-\varepsilon} = \mathcal{O}(n^{1/4-2+0.03+5/3+0.01}) = o(1)$. We deduce that the second term of (2) is dominated by $kn^{-1/3+4\varepsilon}$, and then

$$\frac{\theta_k^{(n+1)} |\mu_k^{(n)}|^2}{\theta_k^{(n)} - \theta_k^{(n+1)}} = 1 + \mathcal{O}(kn^{-\frac{1}{3}+5\varepsilon}).$$

Changing the value of ε appropriately gives the desired result. \square

This lemma is enough for us to estimate the change in $\theta_k^{(n)}$ as n grows, and in particular to find a limit for the renormalized angle. We can now complete the proof of Theorem 1.1.

PROOF OF THEOREM 1.1. The proof proceeds exactly as in [3]. It is sufficient to prove the result for ε equal to the inverse of an integer: hence, it is enough to show the estimate for fixed ε . By symmetry, one can take $k > 0$. We rearrange the equation in Lemma 4.1 to find

$$|\mu_k^{(n)}|^2 = \left(\frac{\theta_k^{(n)}}{\theta_k^{(n+1)}} - 1 \right) (1 + \mathcal{O}(kn^{-\frac{1}{3}+\varepsilon})).$$

Because almost surely, $|\mu_k^{(n)}|^2 = \mathcal{O}(n^{-1+2\varepsilon})$, we get

$$|\mu_k^{(n)}|^2 = \frac{\theta_k^{(n)}}{\theta_k^{(n+1)}} - 1 + \mathcal{O}(kn^{-\frac{4}{3}+3\varepsilon}).$$

Using the asymptotic $\log(1 - \delta) = -\delta + \mathcal{O}(\delta^2)$ for $\delta = o(1)$, we conclude, if ε is small enough,

$$\log \frac{\theta_k^{(n)}}{\theta_k^{(n+1)}} = |\mu_k^{(n)}|^2 + \mathcal{O}(kn^{-\frac{4}{3}+3\varepsilon}).$$

Define the random variable $L_k^{(n)} = \log \theta_k^{(n)} + \sum_{j=1}^{n-1} |\mu_k^{(j)}|^2$; we have just shown $L_k^{(n+1)} - L_k^{(n)} = \mathcal{O}(kn^{-\frac{4}{3}+3\varepsilon})$ so for k fixed, $L_k^{(n)}$ converges to a limit $L_k^{(\infty)}$ almost

surely as $n \rightarrow \infty$, with $|L_k^{(n)} - L_k^{(\infty)}| = \mathcal{O}(kn^{-\frac{1}{3}+3\varepsilon})$. Now,

$$\begin{aligned} \exp L_k^{(n)} &= \theta_k^{(n)} \exp \sum_{j=k}^{n-1} |\mu_k^{(j)}|^2 \\ &= n\theta_k^{(n)} \exp \left(-\log n + \sum_{j=1}^{n-1} \frac{1}{j} + \sum_{j=1}^{n-1} \left(|\mu_k^{(j)}|^2 - \frac{1}{j} \right) \right). \end{aligned}$$

Recall $-\log n + \sum_{j=1}^{n-1} \frac{1}{j} = \gamma + \mathcal{O}(n^{-1})$ where γ is the Euler–Mascheroni constant. Next, we define

$$M_k^{(n)} := \sum_{j=1}^{n-1} \left(|\mu_k^{(j)}|^2 - \frac{1}{j} \right)$$

and observe that each term of the sum is an independent mean-zero random variable. Therefore, for k fixed, $(M_k^{(n)})_{n \geq k}$ is a martingale. We claim that $M_k^{(n)}$ is bounded in L^2 ; in fact,

$$\mathbb{E} \left(|\mu_k^{(n)}|^2 - \frac{1}{n} \right)^2 = \mathcal{O}(n^{-2}),$$

so that

$$\mathbb{E}((M_k^{(\infty)} - M_k^{(n)})^2) = \sum_{j \geq n} \mathbb{E} \left(|\mu_k^{(j)}|^2 - \frac{1}{j} \right)^2 = \mathcal{O}(n^{-1}),$$

where $M_k^{(\infty)}$ is the claimed limit of $M_k^{(n)}$ (this limit exists since $M_k^{(n)}$ is a sum of centered and independent random variables with summable variances). To see this, we observe that $|\mu_k^{(n)}|^2$ is a Beta random variable of parameters 1 and $n-1$, whose variance can be explicitly computed, and is easily checked to be dominated by n^{-2} . Now, by the triangle inequality and Doob's maximal inequality, for q positive integer, $k \leq 2^q$,

$$\begin{aligned} \mathbb{E} \left(\sup_{n \geq 2^q} (M_k^{(\infty)} - M_k^{(n)})^2 \right) &\lesssim \mathbb{E}((M_k^{(\infty)} - M_k^{(2^q)})^2) + \mathbb{E} \left(\sup_{n \geq 2^q} (M_k^{(n)} - M_k^{(2^q)})^2 \right) \\ &\lesssim \mathbb{E}(M_k^{(\infty)} - M_k^{(2^q)})^2 \\ &= \mathcal{O}(2^{-q}). \end{aligned}$$

Hence,

$$\begin{aligned} \mathbb{E} \left[\sup_{2^q \leq n \leq 2^{q+1}} \sup_{k \leq n^{1/4}} (M_k^{(\infty)} - M_k^{(n)})^2 \right] &\leq \mathbb{E} \left[\sup_{k \leq 2^{(q+1)/4}} \sup_{2^q \leq n \leq 2^{q+1}} (M_k^{(\infty)} - M_k^{(n)})^2 \right] \\ &\leq \sum_{k \leq 2^{(q+1)/4}} \mathbb{E} \left[\sup_{2^q \leq n \leq 2^{q+1}} (M_k^{(\infty)} - M_k^{(n)})^2 \right] \\ &\lesssim 2^{(q+1)/4} 2^{-q} = \mathcal{O}(2^{-3q/4}) \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}\left[\sup_{n \geq 2^q} \sup_{k \leq n^{1/4}} (M_k^{(\infty)} - M_k^{(n)})^2\right] &\leq \sum_{r \geq q} \mathbb{E}\left[\sup_{2^r \leq n \leq 2^{r+1}} \sup_{k \leq n^{1/4}} (M_k^{(\infty)} - M_k^{(n)})^2\right] \\ &\lesssim \sum_{r \geq q} 2^{-3r/4} = \mathcal{O}(2^{-3q/4}). \end{aligned}$$

By Markov's inequality, we get

$$\begin{aligned} \mathbb{P}\left(\sup_{n \geq 2^q} \sup_{k \leq n^{1/4}} |M_k^{(\infty)} - M_k^{(n)}| \geq 2^{-q/3}\right) &\leq 2^{2q/3} \mathbb{E}\left(\sup_{n \geq 2^q} \sup_{k \leq n^{1/4}} (M_k^{(\infty)} - M_k^{(n)})^2\right) \\ &= \mathcal{O}(2^{-q/12}), \end{aligned}$$

which, by the Borel–Cantelli lemma, shows that almost surely for some $q_0 \geq 1$, all $q \geq q_0$, $n \geq 2^q$ and $k \leq n^{1/4}$ satisfy $|M_k^{(\infty)} - M_k^{(n)}| \leq 2^{-q/3}$. Hence,

$$|M_k^{(\infty)} - M_k^{(n)}| = \mathcal{O}(n^{-\frac{1}{3}})$$

almost surely. Collecting these estimates and applying them to the equation

$$\exp L_k^{(n)} = n\theta_k^{(n)} \exp(\gamma + \mathcal{O}(n^{-1}) + M_k^{(n)})$$

gives us

$$\exp(L_k^{(\infty)} + \mathcal{O}(kn^{-\frac{1}{3}+3\varepsilon})) = n\theta_k^{(n)} \exp(\gamma + M_k^{(\infty)} + \mathcal{O}(n^{-\frac{1}{3}}))$$

Rearranging,

$$n\theta_k^{(n)} = \exp(L_k^{(\infty)} - M_k^{(\infty)} - \gamma)(1 + \mathcal{O}(kn^{-\frac{1}{3}+3\varepsilon})) =: 2\pi y_k(1 + \mathcal{O}(kn^{-\frac{1}{3}+3\varepsilon})).$$

Now, by [3], $(y_k)_{k \in \mathbb{Z}}$ is a determinantal sine-kernel process, so we have almost surely the estimate $y_k = \mathcal{O}(1 + |k|)$, which proves Theorem 1.1. \square

5. Weak convergence and renormalization of the eigenvectors. We are now ready to show that the eigenfunctions $f_k^{(n)}$ of u_n converge in a suitable sense. We assume that for a given value of ε , the event E from Section 4 holds, which happens almost surely. Recall from Theorem 2.4 that we can choose representatives $f_k^{(n)}$ for the eigenvectors of each u_n in such a way that

$$(h_k^{(n+1)})^{\frac{1}{2}} f_k^{(n+1)} = \sum_{j=1}^n \frac{\mu_j^{(n)}}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} f_j^{(n)} + \frac{v_n - 1}{1 - \lambda_k^{(n+1)}} e_{n+1},$$

where

$$h_k^{(n+1)} = \sum_{j=1}^n \frac{|\mu_j^{(n)}|^2}{|\lambda_j^{(n)} - \lambda_k^{(n+1)}|^2} + \frac{|v_n - 1|^2}{|1 - \lambda_k^{(n+1)}|^2}.$$

For $n = 1$, we adopt the convention $f_1^{(1)} = -e_1$. We deduce, for $n \geq \ell$,

$$\langle f_k^{(n+1)}, e_\ell \rangle = (h_k^{(n+1)})^{-\frac{1}{2}} \sum_{j=1}^n \frac{\mu_j^{(n)}}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} \langle f_j^{(n)}, e_\ell \rangle.$$

The invariance by conjugation of the Haar measures implies that each eigenvector $f_k^{(n+1)}$, multiplied by an independent random phase of modulus 1, is a uniform vector on the complex sphere S^{n+1} . Hence, the scalar product $\langle f_k^{(n+1)}, e_\ell \rangle$ converges to zero in probability. In order to get a limit which is different from zero, we need to consider a suitable normalization. We introduce the following eigenvectors, for $n \geq k$:

$$g_k^{(n)} := D_k^{(n)} f_k^{(n)},$$

where $D_k^{(n)} \in \mathbb{C}$ is the random variable

$$D_k^{(n)} = \prod_{s=k}^{n-1} (h_k^{(s+1)})^{\frac{1}{2}} \frac{\lambda_k^{(s)} - \lambda_k^{(s+1)}}{\mu_k^{(s)}}.$$

We claim that for each renormalized eigenvector $g_k^{(n)}$, the scalar product $\langle g_k^{(n)}, e_\ell \rangle$ converges.

THEOREM 5.1. *For each $k \geq 1$ and $\ell \geq 1$, there exists an increasing sequence $(H_j)_{j \geq 1}$ of events in \mathcal{A} , with probability tending to 1, such that for all $j \geq 1$, $(\mathbb{1}_{H_j} \langle g_k^{(n)}, e_\ell \rangle)_{n \geq k \vee \ell}$ is a martingale with respect to the filtration $(\mathcal{B}_n)_{n \geq k \vee \ell}$, and the conditional expectation of $\mathbb{1}_{H_j} |\langle g_k^{(n)}, e_\ell \rangle|^2$, given \mathcal{A} , is almost surely bounded when n varies.*

REMARK 5.2. We introduce the events H_j in order to avoid to define conditional expectation of variables which are not necessarily integrable or nonnegative.

PROOF. From the equation above,

$$\begin{aligned} \langle g_k^{(n+1)}, e_\ell \rangle &= D_k^{(n+1)} (h_k^{(n+1)})^{-\frac{1}{2}} \sum_{j=1}^n \frac{\mu_j^{(n)}}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} \langle f_j^{(n)}, e_\ell \rangle \\ (3) \quad &= D_k^{(n)} \frac{\lambda_k^{(n)} - \lambda_k^{(n+1)}}{\mu_k^{(n)}} \sum_{j=1}^n \frac{\mu_j^{(n)}}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} \langle f_j^{(n)}, e_\ell \rangle \\ &= \langle g_k^{(n)}, e_\ell \rangle + D_k^{(n)} \frac{\lambda_k^{(n)} - \lambda_k^{(n+1)}}{\mu_k^{(n)}} \sum_{\substack{1 \leq j \leq n \\ j \neq k}} \frac{\mu_j^{(n)}}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} \langle f_j^{(n)}, e_\ell \rangle. \end{aligned}$$

Now, recall from Lemma 3.2 that the sequence of eigenvectors $(f_k^{(n)})_{n \geq k}$ is adapted to the filtration of σ -algebras $(\mathcal{B}_n)_{n \geq k}$. If we decompose

$$\mu_j^{(n)} = \phi_j^{(n)} |\mu_j^{(n)}|$$

then for n fixed, $\{\phi_j^{(n)}\}_{1 \leq j \leq n}$ is a family of i.i.d. random phases uniformly distributed on the unit circle, independent of the σ -algebra \mathcal{B}_n . Indeed, \mathcal{B}_n is the σ -algebra generated by \mathcal{A} and $\{\phi_j^{(m)}\}_{1 \leq m \leq n-1, 1 \leq j \leq m}$, and then, by Lemma 3.1, it is the σ -algebra generated by u_1 , $(v_m)_{m \geq 1}$, $(|\mu_j^{(m)}|)_{m \geq 1, 1 \leq j \leq m}$, and $(\phi_j^{(m)})_{1 \leq m \leq n-1, 1 \leq j \leq m}$. All these variables are independent of $(\phi_j^{(n)})_{1 \leq j \leq n}$, which are i.i.d. uniform on the unit circle, since the vectors $((\mu_j^{(m)})_{1 \leq j \leq m}, v_m)$ are independent, uniform on the unit complex sphere of \mathbb{C}^{m+1} .

Now, since $h_k^{(n+1)}$ is real, we have

$$\Phi_k^{(n)} := \frac{D_k^{(n)}}{|D_k^{(n)}|} = \prod_{s=k}^{n-1} (\phi_k^{(s)})^{-1} \frac{\lambda_k^{(s)} - \lambda_k^{(s+1)}}{|\lambda_k^{(s)} - \lambda_k^{(s+1)}|}$$

for $1 \leq k \leq n$, and we can write

$$\langle g_k^{(n+1)} - g_k^{(n)}, e_\ell \rangle = \Phi_k^{(n)} |D_k^{(n)}| \frac{\lambda_k^{(n)} - \lambda_k^{(n+1)}}{|\mu_k^{(n)}|} \sum_{\substack{1 \leq j \leq n \\ j \neq k}} \frac{\phi_j^{(n)}}{\phi_k^{(n)}} \frac{|\mu_j^{(n)}|}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} \langle f_j^{(n)}, e_\ell \rangle.$$

We would like to compute the conditional expectation of the difference $H_j \langle g_k^{(n+1)} - g_k^{(n)}, e_\ell \rangle$, given the σ -algebra \mathcal{B}_n , for suitable events H_j . We first verify the measurability of each quantity on the right; in particular:

1. $|D_k^{(n)}|$, $\frac{\lambda_k^{(n)} - \lambda_k^{(n+1)}}{|\mu_k^{(n)}|}$, $\frac{|\mu_j^{(n)}|}{\lambda_j^{(n)} - \lambda_k^{(n+1)}}$ are \mathcal{A} -measurable.
2. $\Phi_k^{(n)}$ and $\langle f_j^{(n)}, e_\ell \rangle$ are \mathcal{B}_n -measurable.
3. $\{\phi_j^{(n)} (\phi_k^{(n)})^{-1}\}_{\substack{1 \leq j \leq n \\ j \neq k}}$ are i.i.d. and independent of \mathcal{B}_n .

We also have

$$|\langle g_k^{(n+1)} - g_k^{(n)}, e_\ell \rangle| \leq |D_k^{(n)}| \frac{|\lambda_k^{(n)} - \lambda_k^{(n+1)}|}{|\mu_k^{(n)}|} \sum_{\substack{1 \leq j \leq n \\ j \neq k}} \frac{|\mu_j^{(n)}|}{|\lambda_j^{(n)} - \lambda_k^{(n+1)}|}$$

which gives an \mathcal{A} -measurable bound for the scalar product $\langle g_k^{(n+1)} - g_k^{(n)}, e_\ell \rangle$, which is almost surely finite. Let us denote this bound by X_n , and let us define the event H_j as the intersection of the events $\{X_n \leq R_{n,j}\}$ for $n \geq k \vee \ell$, $R_{n,j}$ being increasing in j and chosen in such a way that $\mathbb{P}[X_n > R_{n,j}] \leq j^{-1} 2^{-n}$, which implies that the probability of H_j tends to 1 when j goes to infinity.

By Fubini's theorem, measurability and the fact that the quantity inside the expectation is uniformly bounded by $R_{n,j}$,

$$\begin{aligned} \mathbb{E}[\mathbb{1}_{H_j} \langle g_k^{(n+1)} - g_k^{(n)}, e_\ell \rangle | \mathcal{B}_n] &= \mathbb{1}_{H_j} \Phi_k^{(n)} |D_k^{(n)}| \frac{\lambda_k^{(n)} - \lambda_k^{(n+1)}}{|\mu_k^{(n)}|} \\ &\quad \times \sum_{\substack{1 \leq j \leq n \\ j \neq k}} \mathbb{E} \left[\frac{\phi_j^{(n)}}{\phi_k^{(n)}} | \mathcal{B}_n \right] \frac{|\mu_j^{(n)}|}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} \langle f_j^{(n)}, e_\ell \rangle. \end{aligned}$$

However, by independence,

$$\mathbb{E} \left[\frac{\phi_j^{(n)}}{\phi_k^{(n)}} | \mathcal{B}_n \right] = \mathbb{E} \left[\frac{\phi_j^{(n)}}{\phi_k^{(n)}} \right] = 0,$$

so $(\mathbb{1}_{H_j} \langle g_k^{(n)}, e_\ell \rangle)_{n \geq k \vee \ell}$ is a $(\mathcal{B}_n)_{n \geq k \vee \ell}$ -martingale.

To check its conditional boundedness in L^2 , given \mathcal{A} , it is sufficient to show

$$\mathbb{E}[\mathbb{1}_{H_j} |\langle g_k^{(k \vee \ell)}, e_\ell \rangle|^2 | \mathcal{A}] + \sum_{n \geq k \vee \ell} \mathbb{E}[\mathbb{1}_{H_j} |\langle g_k^{(n+1)} - g_k^{(n)}, e_\ell \rangle|^2 | \mathcal{A}] < \infty$$

almost surely. In fact, we will prove this inequality without the event H_j (here, the conditional expectations are well defined since the variables are nonnegative). The first term is smaller than or equal to $\|g_k^{(k \vee \ell)}\|^2 = |D_k^{(k \vee \ell)}|^2$, which is \mathcal{A} -measurable and almost surely finite. Hence, it is sufficient to bound the sum. Expanding it gives

$$(4) \quad \mathbb{E}[|\langle g_k^{(n+1)} - g_k^{(n)}, e_\ell \rangle|^2 | \mathcal{A}] = |D_k^{(n)}|^2 \frac{|\lambda_k^{(n)} - \lambda_k^{(n+1)}|^2}{|\mu_k^{(n)}|^2} S,$$

where

$$S = \sum_{\substack{1 \leq i, j \leq n \\ i, j \neq k}} \frac{|\mu_i^{(n)}|}{\lambda_i^{(n)} - \lambda_k^{(n+1)}} \frac{|\mu_j^{(n)}|}{\lambda_j^{(n)} - \lambda_k^{(n+1)}} \mathbb{E}[\overline{\phi_i^{(n)}} \phi_j^{(n)} \langle f_i^{(n)}, e_\ell \rangle \overline{\langle f_j^{(n)}, e_\ell \rangle} | \mathcal{A}].$$

In order to show this equality even if some of the variables involved in S are not integrable, we observe that the equality is true if we multiply all the variables by the indicator of the \mathcal{A} -measurable event

$$|D_k^{(n)}|^2 \frac{|\lambda_k^{(n)} - \lambda_k^{(n+1)}|^2}{|\mu_k^{(n)}|^2} \sum_{\substack{1 \leq i, j \leq n \\ i, j \neq k}} \frac{|\mu_i^{(n)}|}{|\lambda_i^{(n)} - \lambda_k^{(n+1)}|} \frac{|\mu_j^{(n)}|}{|\lambda_j^{(n)} - \lambda_k^{(n+1)}|} \leq R,$$

for any $R > 0$. Then we can let $R \rightarrow \infty$. Now

$$\begin{aligned} \mathbb{E}[\overline{\phi_i^{(n)}} \phi_j^{(n)} \langle f_i^{(n)}, e_\ell \rangle \overline{\langle f_j^{(n)}, e_\ell \rangle} | \mathcal{A}] &= \mathbb{E}[\mathbb{E}[\overline{\phi_i^{(n)}} \phi_j^{(n)} | \mathcal{B}_n] \overline{\langle f_i^{(n)}, e_\ell \rangle} \langle f_j^{(n)}, e_\ell \rangle | \mathcal{A}] \\ &= \delta_{i,j} \mathbb{E}[\overline{\langle f_i^{(n)}, e_\ell \rangle} \langle f_j^{(n)}, e_\ell \rangle | \mathcal{A}], \end{aligned}$$

where $\delta_{i,j}$ is the Kronecker delta. Thus

$$(5) \quad S = \sum_{\substack{1 \leq j \leq n \\ j \neq k}} \frac{|\mu_j^{(n)}|^2}{|\lambda_j^{(n)} - \lambda_k^{(n+1)}|^2} \mathbb{E}[|f_j^{(n)}, e_\ell|^2 | \mathcal{A}].$$

In order to effectively bound this sum, we need a posteriori information from the convergence of the eigenvalues in Section 4. We define the event F_k to be

$$F_k := \{\exists n_0 \geq 1, \forall n \geq n_0, (\theta_{k+1}^{(n)} - \theta_k^{(n)}) \wedge (\theta_k^{(n)} - \theta_{k-1}^{(n)}) \geq n^{-1-\varepsilon}\}.$$

LEMMA 5.3. F_k is \mathcal{A} -measurable and holds with probability one.

PROOF. F_k depends only on the eigenangles (hence eigenvalues) and is therefore clearly \mathcal{A} -measurable. By Theorem 1.1 applied to $k-1, k, k+1$, we see that each of the associated eigenangles satisfies

$$n\theta_j^{(n)} = 2\pi y_j + \mathcal{O}(n^{-\frac{1}{4}}).$$

In particular,

$$\theta_k^{(n)} - \theta_{k-1}^{(n)} = 2\pi n^{-1}(y_k - y_{k-1}) + \mathcal{O}(n^{-\frac{5}{4}}).$$

Since $\{y_j\}_{j \in \mathbb{Z}}$ is a sine-kernel point process $y_k - y_{k-1} \gtrsim 1$ almost surely and similar for $k+1$ and k . \square

Now we assume that F_k holds and estimate the quantities involved in $\mathbb{E}[|g_k^{(n+1)} - g_k^{(n)}, e_\ell|^2 | \mathcal{A}]$. We will begin by estimating $D_k^{(n)}$. By Lemma 4.1, almost surely,

$$|\theta_k^{(n+1)} - \theta_k^{(n)}| = \theta_k^{(n+1)} |\mu_k^{(n)}|^2 (1 + \mathcal{O}(n^{-\frac{1}{3}+\varepsilon})),$$

hence

$$(6) \quad |\lambda_k^{(n+1)} - \lambda_k^{(n)}| = \theta_k^{(n+1)} |\mu_k^{(n)}|^2 (1 + \mathcal{O}(n^{-\frac{1}{3}+\varepsilon})).$$

Similarly,

$$\frac{|v_n - 1|^2}{|1 - \lambda_k^{(n+1)}|^2} = (\theta_k^{(n+1)})^{-2} (1 + \mathcal{O}(n^{-\frac{1}{2}+\varepsilon})).$$

Now, as in the proof of convergence of the eigenangles, let us consider the set J of indices j such that $\theta_k^{(n+1)} - \pi < \theta_j^{(n)} \leq \theta_k^{(n+1)} + \pi$. From the event F_k , we see that for $j \in J \setminus \{k\}$, and n large enough,

$$\begin{aligned} |\theta_j^{(n)} - \theta_k^{(n+1)}| &\geq |\theta_k^{(n)} - \theta_j^{(n)}| - |\theta_k^{(n)} - \theta_k^{(n+1)}| \\ &\geq n^{-1-\varepsilon} - \mathcal{O}(\theta_k^{(n+1)} |\mu_k^{(n)}|^2 (1 + \mathcal{O}(n^{-\frac{1}{3}+\varepsilon}))) \\ &\geq n^{-1-\varepsilon} - \mathcal{O}(n^{-2+\varepsilon}) \gtrsim n^{-1-\varepsilon}, \end{aligned}$$

and then

$$(7) \quad |\lambda_j^{(n)} - \lambda_k^{(n+1)}| \gtrsim n^{-1-\varepsilon},$$

if ε is taken small enough. We can get a stronger estimate for $|j - k|$ sufficiently large. In fact, since E_3 holds, we have

$$(8) \quad |\lambda_j^{(n)} - \lambda_k^{(n+1)}| \gtrsim |j - k| n^{-\frac{5}{3}-\varepsilon}.$$

These two lower bounds let us estimate

$$\begin{aligned} \sum_{\substack{1 \leq j \leq n \\ j \neq k}} \frac{|\mu_j^{(n)}|^2}{|\lambda_j^{(n)} - \lambda_k^{(n+1)}|^2} &\lesssim n^{-1+\varepsilon} \sum_{\substack{j \in J \\ j \neq k}} |\lambda_j^{(n)} - \lambda_k^{(n+1)}|^{-2} \\ &\lesssim n^{-1+\varepsilon} (n^{2+\varepsilon} n^{\frac{2}{3}} + n^{\frac{10}{3}+\varepsilon} n^{-\frac{2}{3}}) \\ &\lesssim n^{\frac{5}{3}+\varepsilon}, \end{aligned}$$

where we split the sum into the interval with $|j - k| \leq n^{\frac{2}{3}}$ and its complement.

Now we can estimate $h_k^{(s+1)}$. In fact,

$$\begin{aligned} h_k^{(s+1)} \frac{|\lambda_k^{(s)} - \lambda_k^{(s+1)}|^2}{|\mu_k^{(s)}|^2} &= 1 + \frac{|\lambda_k^{(s)} - \lambda_k^{(s+1)}|^2}{|\mu_k^{(s)}|^2} \left(\sum_{\substack{1 \leq j \leq n \\ j \neq k}} \frac{|\mu_j^{(s)}|^2}{|\lambda_j^{(s)} - \lambda_k^{(s+1)}|^2} + \frac{|v_s - 1|^2}{|1 - \lambda_k^{(s+1)}|^2} \right) \\ &= 1 + \frac{|\lambda_k^{(s)} - \lambda_k^{(s+1)}|^2}{|\mu_k^{(s)}|^2} (\mathcal{O}(s^{\frac{5}{3}+\varepsilon}) + (\theta_k^{(s+1)})^{-2} (1 + \mathcal{O}(s^{-\frac{1}{2}+\varepsilon}))). \end{aligned}$$

Since almost surely, $\theta_k^{(s+1)} = \mathcal{O}(1/s)$, $s^{\frac{5}{3}+\varepsilon} = \mathcal{O}((\theta_k^{(s+1)})^{-2} s^{-\frac{1}{3}+\varepsilon})$, and then

$$h_k^{(s+1)} \frac{|\lambda_k^{(s)} - \lambda_k^{(s+1)}|^2}{|\mu_k^{(s)}|^2} = 1 + \frac{(\theta_k^{(s+1)})^{-2} |\lambda_k^{(s)} - \lambda_k^{(s+1)}|^2}{|\mu_k^{(s)}|^2} (1 + \mathcal{O}(s^{-\frac{1}{3}+\varepsilon})).$$

Now,

$$\begin{aligned} |\lambda_k^{(s)} - \lambda_k^{(s+1)}|^2 &= |\theta_k^{(s)} - \theta_k^{(s+1)}|^2 (1 + \mathcal{O}(|\theta_k^{(s)} - \theta_k^{(s+1)}|)) \\ &= |\theta_k^{(s)} - \theta_k^{(s+1)}|^2 (1 + \mathcal{O}(s^{-1})). \end{aligned}$$

Using Lemma 4.1, one obtains

$$\begin{aligned} h_k^{(s+1)} \frac{|\lambda_k^{(s)} - \lambda_k^{(s+1)}|^2}{|\mu_k^{(s)}|^2} &= 1 + \frac{(\theta_k^{(s+1)})^{-2} |\theta_k^{(s)} - \theta_k^{(s+1)}|^2}{|\mu_k^{(s)}|^2} (1 + \mathcal{O}(s^{-\frac{1}{3}+\varepsilon})) \\ &= 1 + |\mu_k^{(s)}|^2 (1 + \mathcal{O}(s^{-\frac{1}{3}+\varepsilon})). \end{aligned}$$

Applying the bound on $|\mu_k^{(s)}|^2$ given by E_2 and changing the value of ε gives

$$h_k^{(s+1)} \frac{|\lambda_k^{(s)} - \lambda_k^{(s+1)}|^2}{|\mu_k^{(s)}|^2} = 1 + |\mu_k^{(s)}|^2 + \mathcal{O}(s^{-\frac{4}{3}+\varepsilon}).$$

Thus from the expression

$$|D_k^{(n)}|^2 = \prod_{s=k}^{n-1} h_k^{(s+1)} \frac{|\lambda_k^{(s)} - \lambda_k^{(s+1)}|^2}{|\mu_k^{(s)}|^2},$$

we deduce

$$\begin{aligned} |D_k^{(n)}|^2 &= \prod_{s=k}^{n-1} (1 + |\mu_k^{(s)}|^2 + \mathcal{O}(s^{-\frac{4}{3}+\varepsilon})) \\ &= \exp\left(\sum_{s=k}^{n-1} \frac{1}{s}\right) \exp\left(\sum_{s=k}^{n-1} |\mu_k^{(s)}|^2 - \frac{1}{s}\right) \exp\sum_{s=k}^{n-1} \mathcal{O}(s^{-\frac{4}{3}+\varepsilon}). \end{aligned}$$

As before,

$$\exp\sum_{s=k}^{n-1} \frac{1}{s} = k^{-1} n \exp[\gamma(1 + \mathcal{O}(n^{-1}))],$$

where γ is the Euler–Mascheroni constant,

$$\exp\sum_{s=k}^{n-1} \left(|\mu_k^{(s)}|^2 - \frac{1}{s}\right)$$

is equal to $\exp(M_k^{(n-1)} - M_k^{(k)})$ from Section 4, and the last term converges to a limit N_∞ with error $\mathcal{O}(n^{-\frac{1}{3}+\varepsilon})$.

Thus, we have

$$\begin{aligned} (9) \quad |D_k^{(n)}|^2 &= k^{-1} n \exp(\gamma + M_k^{(\infty)} - M_k^{(k)} + N_\infty) (1 + \mathcal{O}(n^{-\frac{1}{3}+\varepsilon})) \\ &=: D_k n (1 + \mathcal{O}(n^{-\frac{1}{3}+\varepsilon})), \end{aligned}$$

where D_k is a nonzero random variable that depends only on k .

We are now ready to estimate $\mathbb{E}[|\langle g_k^{(n+1)} - g_k^{(n)}, e_\ell \rangle|^2 | \mathcal{A}]$. In fact we have, using (6), (9), the estimate $\theta_k^{(n+1)} = \mathcal{O}(1/n)$ and the bound on $|\mu_k^{(n)}|^2$ given by E_2 ,

$$\begin{aligned} \mathbb{E}[|\langle g_k^{(n+1)} - g_k^{(n)}, e_\ell \rangle|^2 | \mathcal{A}] &= |D_k^{(n)}|^2 \frac{|\lambda_k^{(n)} - \lambda_k^{(n+1)}|^2}{|\mu_k^{(n)}|^2} S \\ &\lesssim n |\mu_k^{(n)}|^2 (\theta_k^{(n+1)})^2 S \\ &\lesssim n^{-2+\varepsilon} S \end{aligned}$$

and using (7) and (8),

$$\begin{aligned} S &= \sum_{\substack{1 \leq j \leq n \\ j \neq k}} \frac{|\mu_j^{(n)}|^2}{|\lambda_j^{(n)} - \lambda_k^{(n+1)}|^2} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^2 | \mathcal{A}] \\ &\lesssim n^{-1+\varepsilon} \sum_{\substack{1 \leq j \leq n \\ j \neq k}} |\lambda_j^{(n)} - \lambda_k^{(n+1)}|^{-2} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^2 | \mathcal{A}] \\ &\lesssim n^{-1+\varepsilon} \sum_{j \in J, 0 < |j-k| < n^{\frac{2}{3}}} n^{2+\varepsilon} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^2 | \mathcal{A}] \\ &\quad + n^{-1+\varepsilon} \sum_{j \in J, |j-k| \geq n^{\frac{2}{3}}} |j-k|^{-2} n^{\frac{10}{3}+\varepsilon} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^2 | \mathcal{A}]. \end{aligned}$$

Therefore, it is now sufficient to prove that for some $\varepsilon > 0$, and almost surely,

$$\begin{aligned} \sum_{n \geq k} n^{-3+\varepsilon} \left(\sum_{j \in J, 0 < |j-k| < n^{\frac{2}{3}}} n^{2+\varepsilon} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^2 | \mathcal{A}] \right. \\ \left. + \sum_{j \in J, |j-k| \geq n^{\frac{2}{3}}} |j-k|^{-2} n^{\frac{10}{3}+\varepsilon} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^2 | \mathcal{A}] \right) < \infty. \end{aligned}$$

It is then sufficient to prove that the expectation of the left-hand side is finite. Since $J \subset [k-n-1, k+n]$, one deduces that it is enough to have

$$\begin{aligned} \sum_{n \geq k} n^{-3+\varepsilon} \left(\sum_{0 < |j-k| < n^{\frac{2}{3}}} n^{2+\varepsilon} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^2] \right. \\ \left. + \sum_{j \in J, n^{\frac{2}{3}} \leq |j-k| \leq n+1} |j-k|^{-2} n^{\frac{10}{3}+\varepsilon} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^2] \right) < \infty. \end{aligned}$$

Now, since $f_j^{(n)}$ is, up to a phase of modulus 1, uniform on the sphere of \mathbb{C}^n , one has

$$\mathbb{E}[|f_j^{(n)}, e_\ell|^2] = 1/n,$$

and then one needs only to check

$$\sum_{n \geq k} n^{-3+\varepsilon} \left(\sum_{0 < |j-k| < n^{\frac{2}{3}}} n^{1+\varepsilon} + \sum_{j \in J, n^{\frac{2}{3}} \leq |j-k| \leq n+1} |j-k|^{-2} n^{\frac{7}{3}+\varepsilon} \right) < \infty,$$

which is easy. \square

The result we have proven has the following consequence.

COROLLARY 5.4. *Almost surely, for all $k \in \mathbb{Z}$ and $\ell \geq 1$, the scalar product $\langle g_k^{(n)}, e_\ell \rangle$ converges to a limit $g_{k,\ell}$ when n goes to infinity.*

PROOF. We can fix k and ℓ . Then, for $j \geq 1$, $R > 0$ we consider the sequence of variables

$$(\mathbb{1}_{H_j, Y_j \leq R} \langle g_k^{(n)}, e_\ell \rangle)_{n \geq k \vee \ell},$$

where Y_j denotes the supremum in n of the conditional expectation of $\mathbb{1}_{H_j} |\langle g_k^{(n)}, e_\ell \rangle|^2$ given \mathcal{A} . Since Y_j is \mathcal{A} -measurable, this sequence is a martingale in the filtration $(\mathcal{B}_n)_{n \geq k \vee \ell}$, and this martingale is bounded in L^2 by construction. Hence, it a.s. converges, and $\langle g_k^{(n)}, e_\ell \rangle$ converges with probability at least $1 - \mathbb{P}[H_j, Y_j \leq R]$. Since we have shown that Y_j is a.s. finite, we let $R \rightarrow \infty$ and deduce a convergence with probability at least $1 - \mathbb{P}[H_j]$, and then we let $j \rightarrow \infty$. \square

For each $k \in \mathbb{Z}$, the infinite sequence $g_k := (g_{k,\ell})_{\ell \geq 1} \in \mathbb{C}^\infty$ can be considered as the weak limit of the eigenvector $g_k^{(n)}$ of u_n , when n goes to infinity. We will study the behavior of a suitable renormalization of g_k in the next section.

6. The law of the coefficients of the limiting eigenvector. In the previous section, we have proven the almost sure convergence of each coordinate of the eigenvectors of u_n , after normalization by a factor $D_k^{(n)}$. Using the estimate (9), we see that almost surely, $|D_k^{(n)}|/\sqrt{n}$ tends to a nonzero random limit when n goes to infinity. It is then more elegant to formulate the result of convergence by taking eigenvectors of norm exactly \sqrt{n} . Moreover, such a normalization provides the distribution of the limiting coordinates of the eigenvectors, as stated in Theorem 1.2 proven just below.

PROOF OF THEOREM 1.2. For fixed $k \in \mathbb{Z}$, let us first show the existence of the sequence $(t_{k,\ell})_{\ell \geq 1}$. By symmetry, one can assume $k \geq 1$: in this case, for all $n \geq k$, there exists $\tau_k^{(n)}$ of modulus 1 such that $f_k^{(n)} = \tau_k^{(n)} v_k^{(n)}$. Hence,

$$|D_k^{(n)}| \langle \Phi_k^{(n)} \tau_k^{(n)} v_k^{(n)}, e_\ell \rangle$$

almost surely converges when n goes to infinity. Now, by the estimate (9), $|D_k^{(n)}|/\sqrt{n}$ converges almost surely to a strictly positive constant. One deduces the existence of $(t_{k,\ell})_{\ell \geq 1}$, by taking

$$\psi_k^{(n)} := \Phi_k^{(n)} \tau_k^{(n)}.$$

Let us now check the uniqueness, by supposing that two sequences $(t_{k,\ell})_{\ell \geq 1}$ and $(t'_{k,\ell})_{\ell \geq 1}$ can be constructed from the same virtual rotation $(u_n)_{n \geq 1}$. In this case, there exist, for all $n \geq 1$, two unit eigenvectors $w_k^{(n)}$ and $w_k^{(n)'}$ corresponding to the same eigenvalue, and such that for all $\ell \geq 1$,

$$(10) \quad \sqrt{n} \langle w_k^{(n)}, e_\ell \rangle \xrightarrow{n \rightarrow \infty} t_{k,\ell}$$

and

$$\sqrt{n} \langle w_k^{(n)'}, e_\ell \rangle \xrightarrow{n \rightarrow \infty} t'_{k,\ell}.$$

Since the eigenvalues are almost surely simple, for all $n \geq 1$, there exists $\chi_k^{(n)} \in \mathbb{U}$ such that $w_k^{(n)'} = \chi_k^{(n)} w_k^{(n)}$, which implies

$$(11) \quad \chi_k^{(n)} \sqrt{n} \langle w_k^{(n)}, e_\ell \rangle \xrightarrow{n \rightarrow \infty} t'_{k,\ell}.$$

By comparing (10) and (11), one deduces that $t'_{k,\ell} = \chi_k t_{k,\ell}$, where $\chi_k \in \mathbb{U}$ denotes the limit of any converging subsequence of $(\chi_k^{(n)})_{n \geq 1}$.

Moreover, let us choose the random vectors $(w_k^{(n)})_{k \in \mathbb{Z}, n \geq 1}$ and the random variables $(t_{k,\ell})_{k \in \mathbb{Z}, \ell \geq 1}$ as measurable functions of $(u_n)_{n \geq 1}$, in such a way that (10) is satisfied almost surely. Let $(\psi_k)_{k \in \mathbb{Z}}$ be i.i.d. random variables, independent of $(u_n)_{n \geq 1}$. For all $n \geq 1$, the invariance by conjugation of the Haar measure on $U(n)$ implies that the family of eigenvectors $(\psi_k w_k^{(n)})_{1 \leq k \leq n}$ of u_n forms a Haar-distributed unitary matrix in $U(n)$. One deduces that if L is a finite set of strictly positive integers, and if K is a finite set of integers, then

$$\sqrt{n} \langle (\psi_k w_k^{(n)})_{k \in K}, e_\ell \rangle_{k \in K, \ell \in L}$$

converges in law to a family of i.i.d. standard complex Gaussian variables. Since this family of variables also converges almost surely, one deduces that the limiting variables $(\psi_k t_{k,\ell})_{k \in K, \ell \in L}$ are i.i.d. standard complex and Gaussian. Since the finite sets K and L can be taken arbitrarily, we are done. \square

From Theorem 1.2, one deduces immediately the following.

COROLLARY 6.1. *The limiting coordinates $g_{k,\ell}$ introduced in Corollary 5.4 are almost surely different from zero.*

PROOF. We know that for a suitable normalization of $t_{k,\ell}$,

$$g_{k,\ell} = \lim_{n \rightarrow \infty} |D_k^{(n)}| \langle \Phi_k^{(n)} f_k^{(n)}, e_\ell \rangle,$$

$$t_{k,\ell} = \lim_{n \rightarrow \infty} \sqrt{n} \langle \Phi_k^{(n)} f_k^{(n)}, e_\ell \rangle,$$

and by (9),

$$\sqrt{D_k} = \lim_{n \rightarrow \infty} |D_k^{(n)}| / \sqrt{n}.$$

Combining these limits gives

$$(12) \quad g_{k,\ell} = \sqrt{D_k} t_{k,\ell},$$

which is almost surely nonzero, since $t_{k,\ell}$ is a standard complex Gaussian variable up to multiplication by an independent uniform random variable on the unit circle. \square

The eigenspaces of u_n , generated by the vectors $f_k^{(n)}$, can also be considered as elements of the projective space $\mathbb{P}^{n-1}(\mathbb{C})$. Moreover, one can define the infinite-dimensional projective space $\mathbb{P}^\infty(\mathbb{C})$, as the space of nonzero infinite sequences of complex numbers, quotiented by scalar multiplication. The convergence of renormalized eigenvectors proven above can be viewed as a convergence of the corresponding points on the projective spaces.

There exists a uniform measure on all these projective spaces, obtained by taking the equivalence class of a sequence of i.i.d. standard complex Gaussian variables. For $n \geq 1$ finite, the uniform measure on $\mathbb{P}^n(\mathbb{C})$ can also be obtained from a uniform point on the sphere in \mathbb{C}^{n+1} . For $m < n \in \mathbb{N} \cup \{\infty\}$, there exists a natural projection $\Pi_{n,m}$ from $\mathbb{P}^n(\mathbb{C})$ to $\mathbb{P}^m(\mathbb{C})$, obtained by taking only the $m+1$ first coordinates of the sequences, and this projection is well defined when these coordinates are not all vanishing: in particular, almost surely under the uniform measure on $\mathbb{P}^n(\mathbb{C})$. Note that the image of this measure by $\Pi_{n,m}$ is the uniform measure on $\mathbb{P}^m(\mathbb{C})$. Moreover, we can define the notion of weak convergence on the projective spaces as follows. Let $(x_n)_{n \geq 1}$ be a sequence such that $x_n \in \mathbb{P}^n(\mathbb{C})$ for all $n \geq 1$ and let $x_\infty \in \mathbb{P}^\infty(\mathbb{C})$. We say that $(x_n)_{n \geq 1}$ weakly converges to x_∞ if and only if the following holds: for all $m \geq 1$ such that the $m+1$ first coordinates of x_∞ are not all vanishing, the projection $\Pi_{n,m}(x_n) \in \mathbb{P}^m(\mathbb{C})$ is well defined for n large enough and tends to $\Pi_{\infty,m}(x_\infty)$ when n goes to infinity. From Theorem 1.2, we can easily deduce the following.

THEOREM 6.2. *Let $(u_n)_{n \geq 1}$ be a virtual rotation, following the Haar measure. For $k \in \mathbb{Z}$ and $n \geq 1$, let $x_k^{(n)} \in \mathbb{P}^{n-1}(\mathbb{C})$ be the eigenspace corresponding to the*

k th smallest nonnegative eigenangle of u_n for $k \geq 1$, and the $(1 - k)$ th largest strictly negative eigenangle of u_n for $k \leq 0$ (this eigenspace is almost surely one-dimensional). Then there almost surely exists some random points $(x_k^{(\infty)})_{k \in \mathbb{Z}}$ in $\mathbb{P}^\infty(\mathbb{C})$ such that for all $k \in \mathbb{Z}$, $x_k^{(n)}$ weakly converges to $x_k^{(\infty)}$ when n goes to infinity. The points $(x_k^{(\infty)})_{k \in \mathbb{Z}}$ are represented by the sequences $(t_{k,\ell})_{\ell \geq 1}$, $k \in \mathbb{Z}$ given above: they are independent and uniform on $\mathbb{P}^\infty(\mathbb{C})$.

APPENDIX A: A PRIORI ESTIMATES FOR UNITARY MATRICES

Let us fix $\varepsilon > 0$. The goal of this section is the proof that the event $E := E_0 \cap E_1 \cap E_2 \cap E_3$ holds almost surely under the Haar measure on the space of virtual isometries, for

$$E_0 = \{\theta_0^{(1)} \neq 0\} \cap \{\forall n \geq 1, v_n \neq 0\} \cap \{\forall n \geq 1, 1 \leq k \leq n, \mu_k^{(n)} \neq 0\},$$

$$E_1 = \{\exists n_0 \geq 1, \forall n \geq n_0, |v_n| \leq n^{-\frac{1}{2}+\varepsilon}\},$$

$$E_2 = \{\exists n_0 \geq 1, \forall n \geq n_0, 1 \leq k \leq n, |\mu_k^{(n)}| \leq n^{-\frac{1}{2}+\varepsilon}\},$$

$$E_3 = \{\exists n_0 \geq 1, \forall n \geq n_0, k \geq 1, n^{-\frac{5}{3}-\varepsilon} \leq \theta_{k+1}^{(n)} - \theta_k^{(n)} \leq n^{-1+\varepsilon}\}.$$

REMARK A.1. In [3], an analogous event was defined: in the present paper, we choose the exponents more carefully to sharpen our results.

We begin by showing that for any fixed basis of \mathbb{C}^n , the coefficients of a uniform random vector on the unit sphere are almost surely $\mathcal{O}(n^{-\frac{1}{2}+\varepsilon})$ for any $\varepsilon > 0$.

LEMMA A.2. Suppose $v_1, \dots, v_n \in \mathbb{C}^n$ is an orthonormal basis and $x \in \mathbb{C}^n$, $\|x\| = 1$ is chosen uniformly from the unit sphere. Then if we write $x = x_1 v_1 + \dots + x_n v_n$, we have bound

$$\mathbb{P}(|x_j|^2 > \delta) = \mathcal{O}(\exp(-\delta n/2)),$$

for all $\delta > 0$ and $j = 1, \dots, n$.

REMARK A.3. We will prove this statement for deterministic vectors $v_1, \dots, v_n \in \mathbb{C}^n$. However, by conditioning, one deduces that the result remains true if the vectors v_1, \dots, v_n are random, as soon as they are independent of x .

PROOF. The random variable $|x_j|^2$ is Beta distributed with parameters 1 and $n - 1$, and then the probability that it is larger than δ is

$$(n - 1) \int_{\delta \wedge 1}^1 (1 - x)^{n-2} dx = (1 - \delta)_+^{n-1} \leq e^{-\delta(n-1)},$$

which shows the result for all $n \geq 2$. For $n = 1$, the probability is 0 for $\delta \geq 1$ and 1 for $\delta \in (0, 1)$, and then smaller than $2e^{-\delta/2}$. \square

From this estimate, we deduce the following bound on the coordinates of the eigenvectors $f_k^{(n)}$.

LEMMA A.4. *Let $\varepsilon > 0$. Then, almost surely, we have*

$$\sup_{1 \leq j, \ell \leq n} |\langle f_j^{(n)}, e_\ell \rangle|^2 = \mathcal{O}(n^{-1+\varepsilon})$$

and

$$\sup_{1 \leq j, \ell \leq n} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^2 | \mathcal{A}] = \mathcal{O}(n^{-1+\varepsilon}),$$

where the implied constant may depend on ε and $(u_m)_{m \geq 1}$.

PROOF. Consider the vector $f_j^{(n)}$ for each fixed j and n . By the invariance by conjugation of the Haar measure on $U(n)$, this eigenvector is, up to multiplication by a complex number of modulus 1, a uniform vector on the unit sphere of \mathbb{C}^n . More precisely, if $\xi \in \mathbb{C}$ is uniform on the unit circle, and independent of $f_j^{(n)}$, then $\xi f_j^{(n)}$ is uniform on the unit sphere. One deduces that for all n, j, ℓ ,

$$\mathbb{P}(|\langle f_j^{(n)}, e_\ell \rangle|^2 > n^{-1+\varepsilon}) = \mathcal{O}(\exp(-n^\varepsilon/2)).$$

Using the Borel–Cantelli lemma gives the first result. Moreover,

$$\begin{aligned} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\varepsilon}] &= \int_0^\infty \mathbb{P}[|\langle f_j^{(n)}, e_\ell \rangle|^2 \geq \delta^{\varepsilon/4}] d\delta \\ &\lesssim \int_0^\infty e^{-n\delta^{\varepsilon/4}/2} d\delta \\ &= \int_0^\infty e^{-z^{\varepsilon/4}/2} d(z/n^{4/\varepsilon}) = \mathcal{O}(n^{-4/\varepsilon}). \end{aligned}$$

We deduce

$$\begin{aligned} \mathbb{P}(\mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\varepsilon} | \mathcal{A}] \geq n^{4-\frac{4}{\varepsilon}}) &\leq n^{\frac{4}{\varepsilon}-4} \mathbb{E}[\mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\varepsilon} | \mathcal{A}]] \\ &= n^{\frac{4}{\varepsilon}-4} \mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\varepsilon}] = \mathcal{O}(n^{-4}). \end{aligned}$$

By the Borel–Cantelli lemma, for all but finitely many $n \geq 1$, $1 \leq j, \ell \leq n$,

$$\mathbb{E}[|\langle f_j^{(n)}, e_\ell \rangle|^{8/\varepsilon} | \mathcal{A}] \leq n^{4-\frac{4}{\varepsilon}}.$$

By the Hölder inequality applied to the conditional expectation, for ε sufficiently small,

$$\mathbb{E}[|f_j^{(n)}, e_\ell|^2 | \mathcal{A}] \leq (\mathbb{E}[|f_j^{(n)}, e_\ell|^{8/\varepsilon} | \mathcal{A}])^{\varepsilon/4} \leq n^{-1+\varepsilon}. \quad \square$$

Another consequence of Lemma A.2 is the following.

PROPOSITION A.5. *The events E_0, E_1, E_2 all hold almost surely.*

PROOF. We apply Lemma A.2 to the decomposition

$$x_{n+1} = \sum_{j=1}^n \mu_j^{(n)} f_j^{(n)} + v_n e_{n+1},$$

which gives

$$\mathbb{P}(|\mu_k^{(n)}|^2 > n^{-1+\varepsilon}) = \mathcal{O}(\exp(-n^\varepsilon/2))$$

so, in particular,

$$\sum_{n \geq 1} \sum_{1 \leq k \leq n} \mathbb{P}(|\mu_k^{(n)}|^2 > n^{-1+\varepsilon}) = \mathcal{O}(1).$$

Therefore, by the Borel–Cantelli lemma, almost surely only a finite number of the events $\{|\mu_k^{(n)}|^2 > n^{-1+\varepsilon}\}$ hold simultaneously. A similar argument controls the coefficients v_n . \square

Before we can control E_3 , we require some estimates on the eigenvalues of a Haar unitary random matrix. Recall that if u_n is distributed according to the Haar measure, then one can define, for $1 \leq p \leq n$, the p -point correlation function $\rho_p^{(n)}$ of the eigenangles, as follows: for any bounded, measurable function ϕ from \mathbb{R}^p to \mathbb{R} ,

$$\begin{aligned} & \mathbb{E} \left[\sum_{1 \leq j_1 \neq \dots \neq j_p \leq n} \phi(\theta_{j_1}^{(n)}, \dots, \theta_{j_p}^{(n)}) \right] \\ &= \int_{[0, 2\pi]^p} \rho_p^{(n)}(t_1, \dots, t_p) \phi(t_1, \dots, t_p) dt_1 \cdots dt_p. \end{aligned}$$

Moreover, if the kernel K is defined by

$$K(t) := \frac{\sin(nt/2)}{2\pi \sin(t/2)}$$

then the p -point correlation function can be given by

$$\rho_p^{(n)}(t_1, \dots, t_n) = \det(K(t_j - t_k))_{j,k=1}^p.$$

Let us first show that the gaps between eigenvalues cannot be asymptotically much larger than average.

LEMMA A.6. *Let $I \subseteq [0, 2\pi)$ be Lebesgue measurable. Then*

$$\mathbb{P}(\text{all of the eigenvalues of } u_n \text{ are in } I) \leq \exp\left(-\frac{|I^c|}{2\pi}n\right).$$

PROOF. We recall the Andreiev–Heine identity, which says that

$$\mathbb{P}(\text{all of the eigenvalues of } u_n \text{ are in } I) = \det M^I,$$

where M^I is an $n \times n$ matrix with entries

$$M_{j,k}^I = \int_I \exp(i(j-k)t) \frac{dt}{2\pi}$$

for j, k between 1 and n . Note that the matrix $(\exp(i(j-k)t))_{j,k=1}^n$ is hermitian and positive, M^I is also; likewise M^{I^c} . Moreover, by computing the entries of $M^I + M^{I^c}$, one checks that this sum is the identity matrix: hence, M^I, M^{I^c} have the same eigenvectors and, if we denote by $(\tau_j)_{1 \leq j \leq n}$ the eigenvalues of M^{I^c} , then $(1 - \tau_j)_{1 \leq j \leq n}$ are the eigenvalues of M^I . The eigenvalues of each matrix must lie in the interval $[0, 1]$, as otherwise one of the eigenvalues of the other matrix would be negative. Now,

$$\det M^I = \prod_{j=1}^n (1 - \tau_j) \leq \exp\left(-\sum_{j=1}^n \tau_j\right) = \exp(-\operatorname{Tr} M^{I^c}) = \exp\left(-\frac{|I^c|}{2\pi}n\right)$$

as was to be shown. \square

Note that the previous lemma applies to all measurable subsets, although we will only need to apply it to intervals.

Next, we control the gaps between eigenvalues from below.

LEMMA A.7. *Suppose $t_1, \dots, t_p \in I$ lie in an interval of length $|I| = \delta \leq 1/n$. Then we have the estimate*

$$\rho_p^{(n)}(t_1, \dots, t_p) = \mathcal{O}_p(\delta^{2p-2} n^{3p-2}).$$

PROOF. We have

$$\rho_p^{(n)}(t_1, \dots, t_p) = \det(K(t_i - t_j))_{i,j=1}^p.$$

The Taylor series for the sine function shows that for $|t| \leq 1/n$,

$$K(t) = \frac{n}{2\pi} \left(1 - \frac{1}{24}(n^2 - 1)t^2 + \mathcal{O}(n^4 t^4)\right).$$

Thus, we have

$$\rho_p^{(n)}(t_1, \dots, t_p) = \frac{n^p}{(2\pi)^p} \det\left(1 - \frac{1}{24}(n^2 - 1)(t_i - t_j)^2 + \mathcal{O}(n^4(t_i - t_j)^4)\right)_{i,j=1}^p.$$

Let A denote the $p \times p$ matrix in the last display, let $\mathbf{1}$ denote the column vector of all ones and let w_j denote the column vector whose i th entry is

$$(w_j)_i = 1 - A_{ij} = \frac{1}{24}(n^2 - 1)(t_i - t_j)^2 + \mathcal{O}(n^4(t_i - t_j)^4).$$

Then by multilinearity and the inclusion-exclusion principle,

$$\det A = \sum_{\sigma \subset [p]} (-1)^{|\sigma|} \det(v_1 \ \cdots \ v_p) \quad \text{where } v_j = \begin{cases} w_j, & j \in \sigma, \\ 1, & \text{otherwise.} \end{cases}$$

Clearly, each term is zero if more than one of the columns is equal to $\mathbf{1}$, so we get

$$\det A = (-1)^{p-1} \sum_{j=1}^p \det M_j + (-1)^p \det M,$$

where M is the matrix with columns w_1, \dots, w_p and M_j is M with the j th column replaced with $\mathbf{1}$. Then in the expansion of each determinant we can bound each term by $\mathcal{O}_p((n^2 \delta^2)^{p-1})$, and the conclusion follows. \square

PROPOSITION A.8. *The event E_3 holds almost surely.*

PROOF. Fix $n \geq 1$. The probability that two adjacent eigenvalues of u_n differ by at least 2δ is bounded above by the probability that one of the parts of the partition

$$(\mathbf{0}, \delta) \cup (\delta, 2\delta) \cup \cdots \cup (\lfloor 2\pi\delta^{-1} \rfloor \delta, \lfloor 2\pi\delta^{-1} + 1 \rfloor \delta)$$

contains no eigenvalue. This, by Lemma A.6, is bounded by

$$\lfloor 2\pi\delta^{-1} + 1 \rfloor \exp(-\delta n/2\pi).$$

Now we let $\delta = n^{-1+\varepsilon}/2$ and apply the Borel–Cantelli lemma to show that at most a finite number of the u_n have gaps larger than $n^{-1+\varepsilon}$.

Next, we see by Lemma A.7 that the probability that two adjacent eigenvalues of u_n differ by at most $\delta \leq 1/n$ is bounded by

$$\iint_{|t_1 - t_2| < \delta} \rho_2^{(n)}(t_1, t_2) dt_1 dt_2 = \mathcal{O}(n^4 \delta^3)$$

which, when we specialize $\delta = n^{-\frac{5}{3}-\varepsilon}$, is $\mathcal{O}(n^{-1-\varepsilon})$; summing over n and applying the Borel–Cantelli lemma shows that these events occur at most a finite number of times as well. \square

APPENDIX B: PROOF OF PROPOSITION 2.1

For all $r \geq 1$, the r -point correlation function of the point process E_n given in Proposition 2.1 is uniformly bounded by 1 and converges pointwise to the correlation function of the determinantal sine-kernel process. The convergence is a classical consequence of the determinantal structure of the eigenvalues of the CUE: see Mehta [10], Section 11.1, for example. The uniform bound is due to the fact that the determinant involved in the correlation functions of the CUE can be written as a Gram determinant, necessarily smaller than or equal to the product of its diagonal entries: for all $\theta_1, \theta_2 \in [0, 2\pi]$,

$$\frac{\sin(n(\theta_2 - \theta_1)/2)}{\sin((\theta_2 - \theta_1)/2)} = \left\langle \sum_{p=0}^{n-1} e^{i\theta_1(p-(n-1)/2)} v_p, \sum_{p=0}^{n-1} e^{i\theta_2(p-(n-1)/2)} v_p \right\rangle,$$

where $(v_p)_{0 \leq p \leq n-1}$ are orthonormal vectors in a complex Hilbert space.

Proposition 2.1 is then a consequence of the following result.

PROPOSITION B.1. *Let E_n denote a point process such that for all $r \geq 1$, its r -point correlation function is well defined, bounded by 1, and converges pointwise to a function $\rho_r^{(\infty)}$ from \mathbb{R}^r to \mathbb{R}_+ . Then there exists a point process E_∞ whose r -point correlation function is $\rho_r^{(\infty)}$ for all $r \geq 1$, and the point process E_n converges to E_∞ in the following sense: for all Borel measurable bounded functions f with compact support from \mathbb{R} to \mathbb{R} ,*

$$\sum_{x \in E_n} f(x) \xrightarrow{n \rightarrow \infty} \sum_{x \in E_\infty} f(x),$$

where the convergence above holds in law.

PROOF. We first note the following identity: for any integer $p \geq 1$,

$$\left(\sum_{x \in E_n} f(x) \right)^p = \sum_{m=1}^{u_p} \sum_{x_1 \neq x_2 \neq \dots \neq x_{r_{p,m}} \in E_n} G_{f,p,m}(x_1, \dots, x_{r_{p,m}}),$$

where u_p depends only on p , $r_{p,m}$ and $m \leq u_p$, and $G_{f,p,m}$ being a measurable, bounded function with compact support from $\mathbb{R}^{r_{p,m}}$ to \mathbb{R} , and depending only on f , p and m . For instance,

$$\begin{aligned} \left(\sum_{x \in E_n} f(x) \right)^3 &= \sum_{x_1 \in E_n} (f(x_1))^3 + 3 \sum_{x_1 \neq x_2 \in E_n} (f(x_1))^2 f(x_2) \\ &\quad + \sum_{x_1 \neq x_2 \neq x_3 \in E_n} f(x_1) f(x_2) f(x_3), \end{aligned}$$

with

$$u_3 = 3, \quad r_{3,1} = 1, \quad r_{3,2} = 2, \quad r_{3,3} = 3,$$

$$G_{f,3,1}(x_1) = (f(x_1))^3,$$

$$G_{f,3,2}(x_1, x_2) = 3(f(x_1))^2 f(x_2),$$

$$G_{f,3,3}(x_1, x_2, x_3) = f(x_1) f(x_2) f(x_3).$$

We can hence write

$$\begin{aligned} & \mathbb{E} \left[\left(\sum_{x \in E_n} f(x) \right)^p \right] \\ &= \sum_{m=1}^{u_p} \int_{\mathbb{R}^{r_{p,m}}} G_{f,p,m}(y_1, \dots, y_{r_{p,m}}) \rho_{r_{p,m}}^{(n)}(y_1, \dots, y_{r_{p,m}}) dy_1 \cdots dy_{r_{p,m}}, \end{aligned}$$

where $\rho_r^{(n)}$ denotes the r -point correlation function of E_n , provided that the above expression converges absolutely, which we now check. Since $G_{f,p,m}$ is measurable, bounded with compact support, we can find $A_{f,p,m} > 0$ such that

$$|G_{f,p,m}(y_1, \dots, y_{r_{p,m}})| \leq A_{f,p,m} \mathbb{1}_{|y_1|, \dots, |y_{r_{p,m}}| \leq A_{f,p,m}}$$

for $y_1, \dots, y_{r_{p,m}} \in \mathbb{R}$. Moreover, we have

$$|\rho_{r_{p,m}}^{(n)}(y_1, \dots, y_{r_{p,m}})| \leq 1$$

by assumption. Consequently, the expression we are dealing with can be bounded from above by

$$\sum_{m=1}^{u_p} \int_{[-A_{f,p,m}, A_{f,p,m}]^{r_{p,m}}} A_{f,p,m} \leq \sum_{m=1}^{u_p} (2A_{f,p,m})^{r_{p,m}+1},$$

which is finite. Moreover, our upper bound is independent of n . Now, we also have by assumption:

$$\rho_{r_{p,m}}^{(n)}(y_1, \dots, y_{r_{p,m}}) \xrightarrow{n \rightarrow \infty} \rho_{r_{p,m}}^{(\infty)}(y_1, \dots, y_{r_{p,m}}),$$

and we can apply the dominated convergence theorem to obtain, for all $p \geq 0$,

$$\mathbb{E}[(X_f^{(n)})^p] \xrightarrow{n \rightarrow \infty} M_{f,p}^{(\infty)},$$

where

$$\begin{aligned} X_f^{(n)} &= \sum_{x \in E_n} f(x), \\ M_{f,p}^{(\infty)} &= \sum_{m=1}^{u_p} \int_{\mathbb{R}^{r_{p,m}}} G_{f,p,m}(y_1, \dots, y_{r_{p,m}}) \rho_{r_{p,m}}^{(\infty)}(y_1, \dots, y_{r_{p,m}}) dy_1 \cdots dy_{r_{p,m}} \end{aligned}$$

if $p \geq 1$, and $M_{f,0}^{(\infty)} = 1$. We also note that

$$\mathbb{E}[|X_f^{(n)}|^p] \leq \mathbb{E}[(X_{|f|}^{(n)})^p] \leq \mathbb{E}\left[\left(\sum_{x \in N} |f(x)|\right)^p\right],$$

where N is a Poisson point process defined on \mathbb{R} with intensity 1 (the last inequality follows from the fact that the correlation functions of E_n are smaller or equal than 1, and hence smaller or equal than the correlation functions of N).

Now for every $\lambda \in \mathbb{R}$, each term of the series

$$\sum_{p \geq 0} \frac{(i\lambda)^p}{p!} \mathbb{E}[(X_f^{(n)})^p]$$

is uniformly dominated in absolute value and independently of n , by the corresponding term in the series

$$\sum_{p \geq 0} \frac{|\lambda|^p}{p!} \mathbb{E}\left[\left(\sum_{x \in N} |f(x)|\right)^p\right] = \mathbb{E}\left[\exp\left(|\lambda| \sum_{x \in N} |f(x)|\right)\right].$$

If we choose $A_f > 0$ in such a way that $|f| \leq A_f$ and such that the support of f is contained in $[-A_f, A_f]$, we have

$$\begin{aligned} \mathbb{E}\left[\exp\left(|\lambda| \sum_{x \in N} |f(x)|\right)\right] &\leq \mathbb{E}[\exp(|\lambda| A_f \text{Card}(N \cap [-A_f, A_f]))] \\ &= \mathbb{E}[e^{|\lambda| A_f Y_{2A_f}}], \end{aligned}$$

Y_{2A_f} standing for a Poisson random variable with parameter $2A_f$. The latter is finite and we can thus apply the dominated convergence theorem to obtain

$$\mathbb{E}[e^{i\lambda X_f^{(n)}}] \xrightarrow{n \rightarrow \infty} \sum_{p \geq 0} \frac{(i\lambda)^p}{p!} M_{f,p}^{(\infty)},$$

the last series in display being absolutely convergent and bounded from above by

$$\begin{aligned} \left|1 - \sum_{p \geq 0} \frac{(i\lambda)^p}{p!} M_{f,p}^{(\infty)}\right| &\leq \sum_{p \geq 1} \frac{|\lambda|^p}{p!} M_{|f|,p}^{(\infty)} \leq \sum_{p \geq 1} \frac{|\lambda|^p}{p!} \sup_{n \geq 1} \mathbb{E}[|X_f^{(n)}|^p] \\ &\leq \sum_{p \geq 1} \frac{|\lambda|^p}{p!} \mathbb{E}\left[\left(\sum_{x \in N} |f(x)|\right)^p\right] \\ &= \mathbb{E}\left[\exp\left(|\lambda| \sum_{x \in N} |f(x)|\right)\right] - 1 \\ &\leq \mathbb{E}[e^{|\lambda| A_f Y_{2A_f}}] - 1 = e^{2A_f(e^{|\lambda| A_f} - 1)} - 1. \end{aligned}$$

Consider now a finite number f_1, f_2, \dots, f_q of measurable and bounded functions with compact support, let $A > 0$ be such that $|f_j| \leq A \mathbb{1}_{[-A, A]}$ for $j \in \{1, \dots, q\}$, and take $\lambda, \lambda_1, \dots, \lambda_q \in \mathbb{R}$. It follows from the definition of $X_f^{(n)}$ that

$$\sum_{j=1}^q \lambda_j X_{f_j}^{(n)} = X_g^{(n)},$$

where

$$g := \sum_{j=1}^q \lambda_j f_j,$$

which implies that

$$\mathbb{E}[e^{i\lambda \sum_{j=1}^q \lambda_j X_{f_j}^{(n)}}] \xrightarrow{n \rightarrow \infty} \sum_{p \geq 0} \frac{(i\lambda)^p}{p!} M_{g,p}^{(\infty)}.$$

Now since g is bounded by $A \sum_{j=1}^q |\lambda_j|$ and since the support of g is included in $[-A, A]$, we have

$$\left| 1 - \sum_{p \geq 0} \frac{(i\lambda)^p}{p!} M_{g,p}^{(\infty)} \right| \leq e^{2A(1 + \sum_{j=1}^q |\lambda_j|)} (e^{|\lambda|A(1 + \sum_{j=1}^q |\lambda_j|)} - 1).$$

If v_1, \dots, v_q are real numbers not all equal to zero, we set

$$\lambda = \sum_{j=1}^q |v_j|, \quad \lambda_j = v_j / \lambda,$$

which implies $\sum_{j=1}^q |\lambda_j| = 1$, and

$$\mathbb{E}[e^{i \sum_{j=1}^q v_j X_{f_j}^{(n)}}] \xrightarrow{n \rightarrow \infty} Q(f_1, \dots, f_q, v_1, \dots, v_q),$$

where

$$|Q(f_1, \dots, f_q, v_1, \dots, v_q) - 1| \leq e^{4A(e^{2|\lambda|A} - 1)} - 1.$$

For fixed f_1, \dots, f_q , the quantity $Q(f_1, \dots, f_q, v_1, \dots, v_q)$ hence tends to 1 when (v_1, \dots, v_q) tends to zero. It follows from Lévy's convergence theorem that the vector $(X_{f_1}^{(n)}, \dots, X_{f_q}^{(n)})$ converges in law, when n goes to infinity, to a random variable with values in \mathbb{R}^q and with characteristic function given by

$$(v_1, \dots, v_q) \mapsto Q(f_1, \dots, f_q, v_1, \dots, v_q).$$

Consequently, we have shown that there exists a family of random variables $(X_f^{(\infty)})_{f \in \mathcal{A}}$ indexed by the set \mathcal{A} of bounded and measurable functions with compact support from \mathbb{R} to \mathbb{R} , satisfying

$$(X_f^{(n)})_{f \in \mathcal{A}} \xrightarrow{n \rightarrow \infty} (X_f^{(\infty)})_{f \in \mathcal{A}},$$

in the sense of finite dimensional distributions. Now for $x \geq 0$, define

$$X^{(n)}(x) = X_{\mathbb{1}_{[0,x]}}^{(n)}, \quad X^{(\infty)}(x) = X_{\mathbb{1}_{[0,x]}}^{(\infty)},$$

and for $x < 0$,

$$X^{(n)}(x) = -X_{\mathbb{1}_{(x,0)}}^{(n)}, \quad X^{(\infty)}(x) = -X_{\mathbb{1}_{(x,0)}}^{(\infty)}.$$

Note that for $y \geq x$, $X^{(n)}(y) - X^{(n)}(x)$ represents the number of points of E_n in the interval $(x, y]$. Moreover, we saw that $(X^{(n)}(x))_{x \in \mathbb{Q}}$ converges in law (in the sense of finite dimensional distributions) to $(X^{(\infty)}(x))_{x \in \mathbb{Q}}$. It follows from Skorokhod's representation theorem that there exist random variables $(Y^{(n)}(x))_{x \in \mathbb{Q}}$ and $(Y^{(\infty)}(x))_{x \in \mathbb{Q}}$, with respectively the same distributions as $(X^{(n)}(x))_{x \in \mathbb{Q}}$ and $(X^{(\infty)}(x))_{x \in \mathbb{Q}}$, such that almost surely, $Y^{(n)}(x)$ converges to $Y^{(\infty)}(x)$ for all $x \in \mathbb{Q}$. By construction, $(Y^{(n)}(x))_{x \in \mathbb{Q}}$ is almost surely integer valued and increasing as a function of x : the same thing holds for $(Y^{(\infty)}(x))_{x \in \mathbb{Q}}$. Moreover, by taking the limits from the right, we can extend $(Y^{(n)}(x))_{x \in \mathbb{Q}}$ et $(Y^{(\infty)}(x))_{x \in \mathbb{Q}}$ to càdlàg functions defined on \mathbb{R} . It is clear that $(Y^{(n)}(x))_{x \in \mathbb{R}}$ has then the same law as $(X^{(n)}(x))_{x \in \mathbb{R}}$, because $(X^{(n)}(x))_{x \in \mathbb{R}}$ is also càdlàg, with the same law when restricted to \mathbb{Q} . We can thus conclude that like for $(X^{(n)}(x))_{x \in \mathbb{R}}$, $(Y^{(n)}(x))_{x \in \mathbb{R}}$ is also the distribution function of some σ -finite measure \mathcal{M}_n , with the same law as the sum of the Dirac measures taken at the points of E_n . Almost surely, for $x \in \mathbb{Q}$, $(Y^{(n)}(x))$ converges to $(Y^{(\infty)}(x))$: hence this convergence also holds at all continuity points of $(Y^{(\infty)}(x))$. Consequently, \mathcal{M}_n converges weakly, in the sense of convergence in law on compact subsets, to a limiting random measure \mathcal{M}_∞ , with distribution function $Y^{(\infty)}$. One can thus write

$$\mathcal{M}_n = \sum_{k \in \mathbb{Z}} \delta_{t_k^{(n)}}, \quad \mathcal{M}_\infty = \sum_{k \in \mathbb{Z}} \delta_{t_k^{(\infty)}},$$

where $\{t_k^{(n)}, k \in \mathbb{Z}\}$ is a set of points with the same distribution as E_n . The weak convergence of \mathcal{M}_n to \mathcal{M}_∞ implies that for $r \geq 0$, F continuous with compact support from \mathbb{R}^r to \mathbb{R} ,

$$\sum_{k_1 \neq k_2 \neq \dots \neq k_r} F(t_{k_1}^{(n)}, \dots, t_{k_r}^{(n)}) \xrightarrow{n \rightarrow \infty} \sum_{k_1 \neq k_2 \neq \dots \neq k_r} F(t_{k_1}^{(\infty)}, \dots, t_{k_r}^{(\infty)}).$$

Indeed the left-hand side can be written as

$$\sum_{m=1}^{u_r} \int_{\mathbb{R}^{s_{r,m}}} H_{F,r,m}(y_1, \dots, y_{s_{r,m}}) d\mathcal{M}_n(y_1) \cdots d\mathcal{M}_n(y_{s_{r,m}}),$$

where u_r depends only on r , $s_{r,m}$ on r and on m and $H_{F,r,m}$ depends on F, r, m , and the right-hand side can be written in similar way with \mathcal{M}_n replaced with \mathcal{M}_∞ .

For instance,

$$\begin{aligned}
 & \sum_{k_1 \neq k_2 \neq k_3} F(t_{k_1}^{(n)}, \dots, t_{k_3}^{(n)}) \\
 &= \int_{\mathbb{R}^3} F(y_1, y_2, y_3) d\mathcal{M}_n(y_1) d\mathcal{M}_n(y_2) d\mathcal{M}_n(y_3) \\
 &\quad - \int_{\mathbb{R}^2} [F(y_1, y_2, y_2) + F(y_2, y_1, y_2) + F(y_1, y_2, y_2)] d\mathcal{M}_n(y_1) d\mathcal{M}_n(y_2) \\
 &\quad + 2 \int_{\mathbb{R}} F(y_1, y_1, y_1) d\mathcal{M}_n(y_1).
 \end{aligned}$$

If we assume that F is positive, it follows from Fatou's lemma that

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{k_1 \neq k_2 \neq \dots \neq k_r} F(t_{k_1}^{(\infty)}, \dots, t_{k_r}^{(\infty)}) \right] \\
 & \leq \liminf_{n \rightarrow \infty} \mathbb{E} \left[\sum_{k_1 \neq k_2 \neq \dots \neq k_r} F(t_{k_1}^{(n)}, \dots, t_{k_r}^{(n)}) \right] \\
 & = \liminf_{n \rightarrow \infty} \int_{\mathbb{R}^r} F(y_1, \dots, y_r) \rho_r^{(n)}(y_1, \dots, y_r) dy_1 \cdots dy_r \\
 & = \int_{\mathbb{R}^r} F(y_1, \dots, y_r) \rho_r^{(\infty)}(y_1, \dots, y_r) dy_1 \cdots dy_r.
 \end{aligned}$$

Reproducing the same computations as in the beginning of our proof yields, for f continuous and positive with compact support, and p a positive integer:

$$N_{f,p}^{(\infty)} := \mathbb{E} \left[\left(\sum_{k \in \mathbb{Z}} f(t_k^{(\infty)}) \right)^p \right] \leq M_{f,p}^{(\infty)}.$$

The bounds that we previously obtained for $M_{f,p}^{(\infty)}$, and which obviously apply to $N_{f,p}^{(\infty)}$ as well, allow us to deduce that for all $\lambda \in \mathbb{R}$,

$$\mathbb{E} \left[\exp \left(i\lambda \sum_{k \in \mathbb{Z}} f(t_k^{(\infty)}) \right) \right] = \sum_{p \geq 0} \frac{(i\lambda)^p}{p!} N_{f,p}^{(\infty)}.$$

Moreover, an application of the dominated convergence theorem yields

$$\begin{aligned}
 \mathbb{E} \left[\exp \left(i\lambda \sum_{k \in \mathbb{Z}} f(t_k^{(\infty)}) \right) \right] &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\exp \left(i\lambda \sum_{k \in \mathbb{Z}} f(t_k^{(n)}) \right) \right] \\
 &= \lim_{n \rightarrow \infty} \mathbb{E} [e^{i\lambda X_f^{(n)}}] = \sum_{p \geq 0} \frac{(i\lambda)^p}{p!} M_{f,p}^{(\infty)}.
 \end{aligned}$$

We can hence conclude that the coefficients of both series in λ are equal, that is, $M_{f,p}^{(\infty)} = N_{f,p}^{(\infty)}$. Going back to the expression of the expansion of the moment of

order p that we gave earlier in the proof, we see that the equality can hold only if

$$\begin{aligned} & \mathbb{E} \left[\sum_{k_1 \neq k_2 \neq \dots \neq k_r} F(t_{k_1}^{(\infty)}, \dots, t_{k_r}^{(\infty)}) \right] \\ &= \int_{\mathbb{R}^r} F(y_1, \dots, y_r) \rho_r^{(\infty)}(y_1, \dots, y_r) dy_1 \cdots dy_r, \end{aligned}$$

for all F, r such that $r = r_{p,m}$, $F = G_{f,p,m}$, with $1 \leq m \leq u_p$. Indeed the left-hand side is always smaller or equal than the right-hand side, and if one of the inequalities were a strict inequality, we would obtain by summing up all terms that $N_{f,p}^{(\infty)} < M_{f,p}^{(\infty)}$. The only term for which $r_{p,m} = p$ gives

$$\begin{aligned} & \mathbb{E} \left[\sum_{k_1 \neq k_2 \neq \dots \neq k_p} F(t_{k_1}^{(\infty)}, \dots, t_{k_p}^{(\infty)}) \right] \\ &= \int_{\mathbb{R}^p} F(y_1, \dots, y_p) \rho_p^{(\infty)}(y_1, \dots, y_p) dy_1 \cdots dy_p, \end{aligned}$$

where

$$F(y_1, \dots, y_p) = f(y_1) \cdots f(y_p).$$

This then extends to all functions F which are measurable, positive and continuous with compact support: indeed there always exists an f which is continuous with compact support on \mathbb{R} such that $F \leq G$, with

$$G(y_1, \dots, y_p) = f(y_1) \cdots f(y_p).$$

Since we have inequalities for both functions F and $G - F$ and an equality for their sum G , we in fact have an equality everywhere.

We see that with a monotone class argument the previous equality then extends to functions F which are measurable, bounded and with compact support. This shows the existence of a point process E_∞ with the same correlation functions as those given in the statement of the proposition, provided we do not exclude a priori point processes with multiple points. More precisely, we take for E_∞ the set of points $t_k^{(\infty)}$ of the support of the measure \mathcal{M}_∞ , taken with their multiplicities. Going back to our earlier computations, we see that for functions f which are measurable, bounded with compact support from \mathbb{R} to \mathbb{R} , and taking into account multiplicities, we have

$$\mathbb{E} \left[\left(\sum_{x \in E_\infty} f(x) \right)^p \right] = M_{f,p}^{(\infty)}$$

and

$$\mathbb{E} \left[\exp \left(i\lambda \sum_{x \in E_\infty} f(x) \right) \right] = \sum_{p \geq 0} \frac{(i\lambda)^p}{p!} M_{f,p}^{(\infty)}.$$

Consequently,

$$\mathbb{E}[e^{i\lambda X_f^{(n)}}] \xrightarrow{n \rightarrow \infty} \mathbb{E}\left[\exp\left(i\lambda \sum_{x \in E_\infty} f(x)\right)\right],$$

which corresponds to the convergence in law stated in the proposition.

It only remains to show that E_∞ does not have multiple points. Indeed, if E_∞ is the set of points $(t_k^{(\infty)})_{k \in \mathbb{Z}}$, taken with multiplicities, then for any measurable bounded function F with compact support from \mathbb{R}^2 in \mathbb{R} ,

$$\mathbb{E}\left(\sum_{k_1 \neq k_2} F(t_{k_1}^{(\infty)}, t_{k_2}^{(\infty)})\right) = \int_{\mathbb{R}^2} F(y_1, y_2) \rho_2^{(\infty)}(y_1, y_2) dy_1 dy_2.$$

Taking $F(y_1, y_2) = \mathbb{1}_{y_1=y_2}$ above yields

$$\mathbb{E}[\text{Card}\{(k_1, k_2) \in \mathbb{Z}^2, k_1 \neq k_2, t_{k_1}^{(\infty)} = t_{k_2}^{(\infty)}\}] = 0,$$

which shows that E_∞ does almost surely not have multiple points. \square

REFERENCES

- [1] BORODIN, A. and OLSHANSKI, G. (2001). Infinite random matrices and ergodic measures. *Comm. Math. Phys.* **223** 87–123. [MR1860761](#)
- [2] BOURGADE, P., HUGHES, C. P., NIKEGHBALI, A. and YOR, M. (2008). The characteristic polynomial of a random unitary matrix: A probabilistic approach. *Duke Math. J.* **145** 45–69. [MR2451289](#)
- [3] BOURGADE, P., NAJNUDEL, J. and NIKEGHBALI, A. (2013). A unitary extension of virtual permutations. *Int. Math. Res. Not. IMRN* **2013** 4101–4134. [MR3106884](#)
- [4] CHHAIBI, R., NAJNUDEL, J. and NIKEGHBALI, A. (2017). The circular unitary ensemble and the Riemann zeta function: The microscopic landscape and a new approach to ratios. *Invent. Math.* **207** 23–113. [MR3592756](#)
- [5] JACOD, J., KOWALSKI, E. and NIKEGHBALI, A. (2011). Mod-Gaussian convergence: New limit theorems in probability and number theory. *Forum Math.* **23** 835–873. [MR2820392](#)
- [6] KATZ, N. M. and SARNAK, P. (1999). *Random Matrices, Frobenius Eigenvalues, and Monodromy*. American Mathematical Society Colloquium Publications **45**. Amer. Math. Soc., Providence, RI. [MR1659828](#)
- [7] KEATING, J. P. and SNAITH, N. C. (2000). Random matrix theory and $\zeta(1/2 + it)$. *Comm. Math. Phys.* **214** 57–89. [MR1794265](#)
- [8] KEROV, S., OLSHANSKI, G. and VERSHIK, A. (1993). Harmonic analysis on the infinite symmetric group. A deformation of the regular representation. *C. R. Acad. Sci. Paris Sér. I Math.* **316** 773–778. [MR1218259](#)
- [9] KOWALSKI, E. and NIKEGHBALI, A. (2010). Mod-Poisson convergence in probability and number theory. *Int. Math. Res. Not. IMRN* **2010** 3549–3587. [MR2725505](#)
- [10] MEHTA, M. L. (2004). *Random Matrices*, 3rd ed. *Pure and Applied Mathematics (Amsterdam)* **142**. Elsevier/Academic Press, Amsterdam. [MR2129906](#)
- [11] NERETIN, Y. A. (2002). Hua-type integrals over unitary groups and over projective limits of unitary groups. *Duke Math. J.* **114** 239–266. [MR1920189](#)

K. MAPLES
A. NIKEGHBALI
INSTITUT FÜR MATHEMATIK
UNIVERSITÄT ZÜRICH
WINTERTHURERSTRASSE 190
CH-8057 ZÜRICH
SWITZERLAND
E-MAIL: kenneth.maples@math.uzh.ch
ashkan.nikeghbali@uzh.ch

J. NAJNUDEL
DEPARTMENT OF MATHEMATICAL SCIENCES
4199 FRENCH HALL WEST
UNIVERSITY OF CINCINNATI
2815 COMMONS WAY
CINCINNATI, OHIO 45221-0025
USA
E-MAIL: joseph.najnudel@uc.edu